Pergamon

PII: S0042-6989(96)00220-9

Vision Res., Vol. 37, No. 9, pp. 1217–1241, 1997

Printed in Great Britain
0042-6989/97 $17.00 + 0.00

Curve Detection in a Noisy Image

ZYGMUNT PIZLO,*‡ MONIKA SALACH-GOLYSKA,* AZRIEL ROSENFELD†

Received 29 December 1995; in revised form 6 June 1996

In this paper we propose a new theory of the Gestalt law of *good continuation*. In this theory perceptual processes are modeled by an exponential pyramid algorithm. To test the new theory we performed three experiments. The subject's task was to detect a target (a set of dots arranged along a straight or curved line) among background dots. Detectability was high when: (a) the target was long; (b) the density of target dots relative to the density of background dots was large; (c) the local change of angle was small along the entire line; (d) local properties of the target were known to the subject. These results are consistent with our new model and they contradict prior models. © 1997 Elsevier Science Ltd. All rights reserved.

Figure–ground segregation Local vs global analysis

INTRODUCTION

The goal of visual perception is to provide the observer with visual information about the three-dimensional environment so that the observer can recognize objects, manipulate them, and navigate in the environment. There have been many theories and models that attempted to describe and explain three-dimensional visual perception (e.g. Braunstein, 1976; Cutting, 1986; Gibson, 1950, 1979; Johansson, 1977; Marr, 1982; Pizlo, 1994; Rock, 1983; Shepard & Cooper, 1982; Zusne, 1970). It is clear, however, that before any three-dimensional object can be reconstructed or recognized, the visual system must first "decide" whether the retina contains an image of any object at all. If it does, the next question is which parts of the retinal image correspond to this object. The phenomenon where a region on the retina representing a given object and the boundary of this region are determined is called *figure–ground segregation*. The importance of this phenomenon was recognized quite early by Gestalt Psychologists (see, for example, Koffka, 1935).

Solving the figure–ground segregation problem may not be easy if the scene is cluttered and the object occluded. This is illustrated in Fig. 1, which shows a house behind trees and bushes. But solving the figure–ground segregation problem is difficult also (or even primarily) because a given retinal image does not uniquely determine the three-dimensional scene: there is always an infinite number of possible interpretations of the regions and contours in the retinal image. However, despite this ambiguity inherent in figure–ground segregation, the observer's percept is usually unambiguous. Figure 2 shows an example. This figure contains a collection of individual dots and there are many possible interpretations of this figure. For example, one could perceive this figure as a set of individual dots unrelated to one another. Alternatively, the observer could *group* the dots, which can lead to the percept of two curved lines. Note, however, that there are at least two different groupings possible because it is unclear whether this figure represents two smooth lines that give rise to an "X" intersection or whether there are two lines, one being a "V" and the other being an inverted "V", meeting at their corners. Clearly, this figure has many possible interpretations. Phenomenologically, however, we tend to see two smooth lines forming an "X" intersection. How does the visual system arrive at this particular interpretation? Why is this interpretation preferred, over many others that are geometrically equivalent?

According to Gestalt Psychologists, to decide among many possible interpretations, the visual system uses some rules or laws. These laws determine how the image is organized before it becomes a percept. Hence, these laws are called *laws of organization*. The fundamental law of organization in Gestalt Psychology is called the *Prägnanz* (or *simplicity*) *principle*. According to this principle, the interpretation that is preferred by the observer is the simplest one that can be derived from a given retinal image. Thus, in the case shown in Fig. 2, the observer perceives two smooth lines, rather than two lines each having a V shape, because in the former case the lines are closer to straight lines, and a straight line is simpler than a V-shaped line. The simplicity principle, when applied to cases similar to that in Fig. 2, is conventionally called the *law of good continuation*. The

*Department of Psychological Sciences, Purdue University, West Lafayette, IN 47907-1364, U.S.A.
†Center for Automation Research, University of Maryland, College Park, MD 20742-3275, U.S.A.
‡To whom all correspondence should be addressed [*Email* pizlo@psych.purdue.edu].

FIGURE 1. A natural scene. This image illustrates the fact that figure–ground segregation is a difficult perceptual problem.

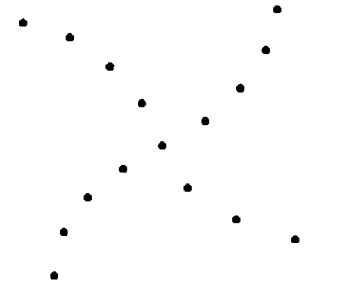

FIGURE 2. Illustration of the law of good continuation. Despite the fact that this set of dots determines many possible interpretations, it usually gives rise to the percept of two smooth lines forming an X intersection.

name of this law, *good continuation*, is related to the perceptual tendency to prefer a smooth line when resolving the direction of a given line at an intersection with another line. This law was described first by Wertheimer (1923/1958). He provided a number of examples illustrating this law and conjectured that good continuation refers to "an appropriateness of the curve, an inner belongingness, a good whole or good configuration". Because of its vagueness, this conjecture did not receive much attention and subsequent theories of good continuation were based on the concept of simplicity, which due to its mathematical tractability could easily lead to testable theories (interestingly, such a formulation of good continuation is not consistent with Wertheimer's understanding of this perceptual phenomenon: he speculated that good continuation "does not imply a mathematical simplicity").

The simplicity principle is not the only principle that can be used to explain the phenomenon of figure–ground segregation. An alternative principle is called the *likelihood principle*. According to this principle our percept follows the interpretation which is the most likely, or the most probable one, that can be derived from a given retinal image. Again, consider the example in Fig. 2. Assume that this figure represents the observer's retinal image produced by two lines in three-dimensions. This retinal image could have been produced by two V-shaped lines in three-dimensions only if the observer viewed these lines from one particular viewing direction, so that the images of the corners of the two lines happened to touch one another. Clearly, such a situation is quite unlikely in everyday life and it happens with probability close to zero (this is called a degenerate view). On the other hand, two smooth lines in three-dimensions can give rise to an X intersection on the retina for a wide range of viewing directions. It is obvious that this can happen with probability greater than zero and, as a result, this situation is more likely as compared to the case with two V-shaped lines. Thus, there have been two different ways of explaining figure–ground segregation, one based on the Prägnanz or simplicity principle of Gestalt and the other based on the likelihood principle of the empiristic school of psychology.

We will begin with a review of prior research, in psychology and in computer vision, on the law of good continuation. Next, we will provide a new theory of this law based on a new definition of smoothness rooted in the likelihood principle. After the smoothness is defined, we will present a stimulation model, based on an exponential pyramid algorithm, of the perceptual mechanism underlying this law. Then we will report the results of three psychophysical experiments that tested the new definition of smoothness as well as the new model.

PRIOR RESEARCH

In this review we will classify the theories and the experimental results with respect to such features as local vs global processing, and bottom-up (data based) vs top-down (model based) processing. Also, we will try to identify whether a given theory or experimental result is consistent with the simplicity principle of Gestalt or the likelihood principle of empiristic psychology.

Traditionally, figure–ground segregation was assumed to be the first stage of perceptual processing, which was a prerequisite for further stages such as object reconstruction and recognition. This assumption implied that in order to ensure that recognition and reconstruction of objects is performed within reasonably short times, this first stage of visual processing had to be fast. Next, since figure–ground segregation by itself was not assumed to lead to elaborated percepts of three-dimensional objects and scenes, it seemed natural to assume that figure–ground segregation involves relatively simple mechanisms. These apparently obvious assumptions, namely, that figure–ground segregation had to be fast and simple, restricted the class of models that have been considered in the past as possible models of figure–ground segregation. We will show, however, that these restrictions might have been too severe and that using them led to inadequate models of figure–ground segregation.

Uttal's (1975) study was one of the first quantitative approaches to curve detection in the presence of noise. Uttal used dots as stimuli. The background was a random dot pattern and a target was some regular arrangement of

dots (e.g. a set of dots along a straight or curved line). If the observer is presented with such a stimulus, the observer can often easily detect the presence of the target. Note that in such dotted stimuli, both the target and the background consist of the same elements, dots, whose only property is position. This allowed Uttal to investigate the role of perceptual organization itself, simply by changing the arrangement of the dots, without changing the physical properties of the individual dots, like intensity, duration or blur.

Uttal tested the effects of various parameters like the density of dots in the target and in the background and the shape of the target, on the target's detectability. The main results of Uttal's experiments can be summarized as follows: the easiest target to detect was a straight line. When the target was a curved line, performance was worse. More precisely, the higher the curvature (i.e. the greater the departure from straightness), the worse was the performance. This result supported the Gestalt simplicity principle. Next, the detectability of the target was higher when:

1. The density of dots in the target was greater;
2. The target was longer; and
3. The density of dots in the background was lower.

If the spacing of dots in the target was irregular, or the positions of the dots in the target were randomly changed so that the target was not a smooth line, detectability was lower.

To account for these results Uttal proposed an autocorrelation model. This model analyzed the entire stimulus globally and this analysis was purely bottom-up (data based). However, the purely global nature of the autocorrelation model is a problem because the output of the model carries little or no information about the target itself (Caelli *et al.*, 1978). Furthermore, if the background is highly regular, for example if it consists of a set of short vertical line segments, and the target is an equally short oblique line segment, the autocorrelation model will not detect the target at all. But the human observer can still quite easily detect a target under such conditions [this phenomenon is called *pop-out*, Treisman & Gelade (1980)]. To overcome some of these problems, other models have been proposed [but see Ben-Av & Sagi (1995) for a recent revival of an autocorrelation model].

van Oeffelen and Vos (1983) proposed a model where processing began with locally convolving the image with a Gaussian distribution function. Then, grouping was determined by finding sets of dots that were located within a single iso-density contour. This operation represented the *proximity law* of Gestalt. The good continuation principle was incorporated into this model by using an elliptical Gaussian distribution with orientation close to the orientation of the line to be found (Smits *et al.*, 1986). The problem with this model was that the orientation of this line had to be known in advance.

To generalize this model to the case of a straight line with unknown orientation as well as to curved lines, Smits and Vos (1986) introduced a connectivity measure. In this model the image was convolved with a circular Gaussian distribution and then the principal directions of curvature of the resulting surface at saddle points were determined. One of these directions was parallel to the line connecting the two dots and the other was perpendicular to this direction. Thus, these directions allowed determination of the orientation of this line. This connectivity measure, along with proximity (measured by the average distance between neighboring dots) and good continuation (measured by the standard deviation of the orientations between adjacent lines), were assumed to affect the saliency measure of a curve.

This approach in which the local analysis involved pairs of dots, has been more recently generalized by including a search for triplets of dots that are approximately collinear (Vos & Helsper, 1991). This new model can detect approximately collinear sets of dots and it may lead to detection of polygonal figures. Vos and Helsper pointed out, however, that this model is only a first step in modeling the phenomenon of good continuation and that an adequate model should be general enough to include the case of curved patterns. Furthermore, they conjectured that a psychologically plausible model cannot be purely bottom-up (data based), as are all current and prior models. Instead, it has to allow for some top-down constraints (hypotheses).

An approach which also involved computing saliency of curves was described by Sha'ashua and Ullman (1988). In their algorithm the saliency of a curve involved concepts of good continuation and proximity. The saliency was greater if the curve was longer, smoother and with fewer gaps (occlusions). The smoothness itself was measured by curvature or curvature variation. Despite the fact that this algorithm can detect curves in similar cases to those where a human observer detects them, this algorithm is not a plausible model of the human perceptual mechanisms underlying curve detection. First, in their algorithm the saliency is accumulated in a serial way along the entire curve. This makes the processing time proportional to the length of the curve. But it is known that the time of perceptual processing of a line is insensitive to the line's length (Pizlo *et al.*, 1995). Another problem is related to using curvature as a measure of smoothness. Although curvature is a measure of departure from straightness curvature is scale dependent (for example, circles having different diameters have different curvatures). As a result, changing size on the retina affects curvatures proportionally, which then should change the perceptual saliency of curves in the scene. This is a problem because in everyday life the retinal sizes of objects change very often when the distances of the objects from the observer's eye change, but the percepts of the contours of the objects do not seem to change [see also Alter & Basri (1996) for their analysis of mathematical aspects related to the scale dependence of Sha'ashua and Ullman's algorithm].

Finally, we will describe briefly an approach based, in part, on the anatomical and physiological properties of the visual cortex (Zucker, 1985; Dobbins *et al.*, 1989).

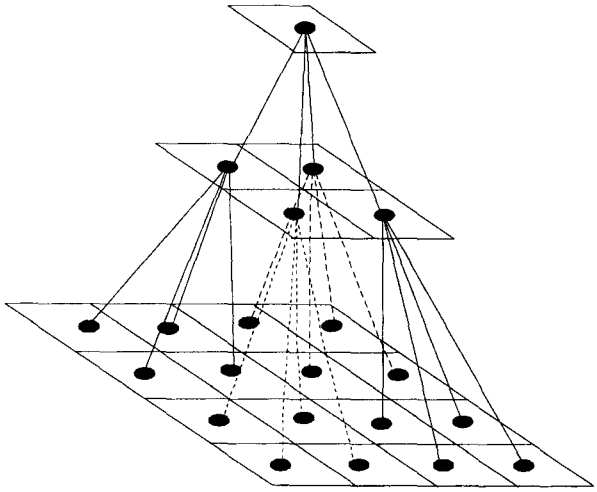

FIGURE 3. Schematic illustration of a non-overlapped exponential pyramid.

Zucker and his colleagues showed that parts of a straight or curved line can be detected by detectors that are sensitive to the orientation and curvature of a line, and whose spatial properties resemble the receptive fields of simple, complex, and hypercomplex cells. They conjectured that after parts of a line are detected, they can be integrated in a curve synthesis stage.

It is seen from our review that prior theories shared two aspects: they all used bottom-up processing and they all involved the simplicity principle. The fact that the prior models did not involve any top-down effects (e.g. familiarity with the target to be detected) allowed for relatively fast processing. However, it is known that human figure–ground segregation can be affected by familiarity with the object to be detected (Leeper, 1935). Therefore, a psychologically plausible model of figure–ground segregation should allow for such effects. Below, we describe a class of algorithms called pyramids whose properties make them a possible class of models of figure–ground segregation. Specifically, they allow performing both local and global operations as well as applying bottom-up and top-down analysis.

Pyramid algorithms

General features. A pyramid consists of a stack of layers containing processing units (nodes) (see Fig. 3 for a schematic illustration of a pyramid). Processing begins at the bottom layer (number 0) which has the largest number of nodes (N). Other layers have fewer nodes: the higher the layer, the smaller the number of nodes. More exactly, the number of nodes in layer k is equal to N/b^k, where b is called a reduction ratio. Because this reduction is exponential, such an architecture is called an exponential pyramid. Each node in the pyramid receives information from a limited region of the image, called a receptive field. The receptive fields of nodes on higher levels are larger. This increase of the receptive field sizes is accomplished by projecting information from several "children" nodes in a lower layer to one "parent" node in a higher layer.

A pyramid has three general properties. First, it allows simultaneous (parallel) processing of different parts of the image. Second, it contains a number of representations of the image, different representations having different spatial scales. Third, it allows for integration of the different representations in either of two ways: fine-to-coarse (bottom-up) or coarse-to-fine (top-down). It is worth pointing out that the traditional distinction between global and local analysis (we used this distinction in the previous section to classify different theories and models) does not really exist in a pyramid. Properties that are local in the higher layers of the pyramid are at the same time global in the lower layers of the pyramid. As a result, pyramids offer a natural solution to the traditional controversy of whether to perform local or global analysis; namely, both types of processing are performed simultaneously.

Exponential pyramid algorithms have been used during the last 20 yr in computer vision to solve a wide range of "early vision" problems like image segmentation and feature extraction [see Tanimoto and Pavlidis (1975) for an early publication on pyramids, and Jolion and Rosenfeld (1994) for a recent review of pyramids]. Interestingly, however, pyramids have not been used very often as models of human vision [Pizlo *et al.* (1995) is one of the few exceptions]. This fact is surprising because on the one hand the pyramid's organization and functioning are similar to the known anatomical and physiological properties of the visual system, and on the other hand, a pyramid can account for a wide range of perceptual observations and experiments. Consider first the anatomy and physiology of the visual system.

Biological plausibility of pyramid models. The human visual system processes visual information in a highly parallel way. Simultaneous activations of different receptors in the retina (there are about 120×10^6 receptors in the retina of each eye) are passed through the optic nerve, which contains 10^6 nerve fibers, to the lateral geniculate nucleus and then to the primary visual cortex (area V1). Nerve fibers that are projected from different parts of the retina terminate at different parts of area V1. This allows simultaneous processing of the visual information. The primary visual cortex is a topographical map of the retina, i.e. neighboring parts of the retina are represented by neighboring parts of V1. Area V1 projects to area V2, which then projects to other areas representing higher stages of visual processing. Thus, the visual system has, to some extent, a hierarchical organization. The separate areas in the visual cortex (V1, V2, V4 etc) are well defined and the numbers reflect the order of processing. Although no area has connections to only one other area, the connections are not random, but instead form quite clear patterns. As a result, processing of the retinal image is done in stages, where the cells at later stages of processing selectively respond to more complex stimulation than the cells at earlier stages of processing. Finally, the sizes of the receptive fields in the visual cortex are different in different areas (Zeki, 1993). The receptive fields of cells in area V1 are the smallest. These

cells project to area V2 in such a way that several cells in V1 converge to one cell in V2. Such connections make the receptive fields of the cells in V2 larger than their counterparts in V1. The cells from area V2 have connections to area V4 (from the thin strips in V2) or V5 (from the thick strips in V2). Again, several cells project to one cell in the higher area, and thus the receptive fields are still larger in areas V4 and V5. So, the higher in the processing hierarchy the cell is, the larger is its receptive field. All these anatomical features of the human visual system: parallel processing, preserving the topographical map of the retina, hierarchical organization, and increasing receptive field sizes at successive stages of processing, are consistent with the organization of the exponential pyramid.

Psychological plausibility of pyramid models. Now, we briefly review psychophysical results that are consistent with the pyramid's architecture. Treisman and Gelade (1980) discovered conditions under which a target can easily be detected among distractors without direction of attention to any particular part of the scene (the "pop-out" phenomenon). Similarly, Julesz (1962) and Beck (1966) determined conditions under which texture segregation occurs effortlessly and pre-attentatively. These two groups of phenomena suggest that the visual system is analyzing the entire scene in a parallel fashion. Next, consider the multiresolution property of the visual system. This property was first demonstrated by Campbell and Robson (1968). They postulated that the human visual system has distinct channels, each of which shows the greatest response to a certain spatial frequency. This work on spatial properties of the visual system has been generalized by Watt (1987) in his model of spatio-temporal integration of visual information. He tested discrimination sensitivity for length, orientation, curvature and stereoscopic depth for stimuli with various sizes and for various exposure durations. He found that the discrimination sensitivity increased with the exposure duration of the stimuli. Watt explained his results by invoking spatial frequency channels and assuming that at the onset of the stimulus only the coarsest channel (resolution) is used. Then, this channel is "switched off" and a finer channel is used, and so on, until the channel with the finest resolution remains. Clearly, this model is similar to (although not identical with) coarse-to-fine processing in the pyramid. This work of Watt was elaborated by Pizlo *et al.* (1995). They pointed out that for a wide range of lengths in Watt's (1987) study, the Weber fraction (the ratio of the discrimination threshold for length to the length itself) was constant [see also Burbeck & Yap (1990) for a similar result]. This result indirectly implied that the speed of length processing in the visual system does not depend on the length itself, but instead on the required precision of length judgment. Pizlo *et al.* (1995) showed that this result can be modeled by an exponential pyramid algorithm and that such a model can better account for a wider range of perceptual phenomena of size perception and mental size transformation than other models.

To summarize, the architecture of the pyramid is a plausible model of the anatomy of the human visual system and the known computational properties of the exponential pyramid seem to agree with results of psychophysical experiments. Therefore, we believe that the exponential pyramid is a possible (perhaps even a plausible) model of the perceptual mechanisms underlying human vision*. In this paper we use this model to study and explain the law of good continuation, which is an element of the figure–ground segregation phenomenon.

In this study we performed three experiments. First we describe Experiment 1 in which we tested the detectability of a set of collinear dots among background dots. Next, we present our exponential pyramid model of curve detection along with a new definition of smoothness. This model was then used to perform simulation experiments under the same conditions as those used in Experiment 1. Then we report Experiment 2 which further tests our model and the new definition of smoothness. In this experiment a number of different curves were used as targets: straight lines, circular arcs and arbitrarily shaped curves. Again, psychophysical results were compared to simulation results from our exponential pyramid model. Finally, our Experiment 3 tests the roles of familiarity of

*One of the reviewers indicated that a pyramid is a discrete precursor of a continuous scale space model (Koenderink, 1984). In this model a given image is represented by a one parameter family of derived images where resolution is a (continuous) parameter. It has to be pointed out that the scale space model is a continuous equivalent of only one special case of an exponential pyramid, called a multiresolution pyramid. In a multiresolution pyramid (and in the scale space model) the only spatial operation which is applied across levels of scale is blurring (which corresponds to computing a weighted average) and this operation is applied to only one feature: intensity. Also, the representation of a stimulus in a multiresolution pyramid (and in the scale space model) is derived in a purely bottom-up fashion. In a general case of an exponential pyramid, on the other hand, one can perform a variety of operations, like computing higher order statistics (e.g. variance) and detecting properties of a frequency distribution such as bimodality or outliers (Rosenfeld, 1990). These operations can be applied to a number of features, and the features themselves can be either continuous (e.g. hue, size, orientation, curvature) or discrete variables (texture elements). Finally, an exponential pyramid allows changing the parameters of the local operations, allowing for top-down (model based) effects. It is possible that some of the capabilities of exponential pyramids could be modeled by continuous models (similar to the scale space model), but we do not think that all discrete pyramids have continuous counterparts. We want to point out, however, that despite the clear theoretical distinction between discrete and continuous models, this distinction may be less important on a computational level. Namely, if the reduction ratio is close to one (as is the case in "fractional pyramids"—Burt, 1981), the discrete pyramid becomes computationally very similar to a continuous model.The question remains of which of the two types of models, discrete or continuous, is psychologically and biologically more plausible. Existing perceptual results, including "pop-out" phenomena, texture segregation and top-down effects, are more consistent with the exponential (discrete) pyramid models, since those models, but not the scale space model, can account for the perceptual results. Similarly, on a biological level, the discrete nature of the nervous system seems to be more consistent with discrete, rather than continuous, models.

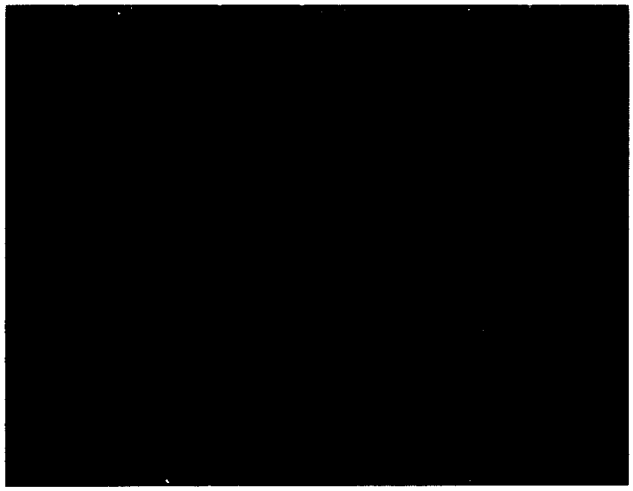

FIGURE 4. A sample stimulus from Experiment 1. The target is the set of collinear dots in the lower right corner. In the experiment, the position and orientation of the target on the screen, as well as its length (or density), were random.

the shape and orientation of the target in its detectability. The last section provides a summary and conclusions.

EXPERIMENT 1: THE EFFECT OF THE DENSITY AND LENGTH OF A TARGET ON ITS DETECTABILITY

In the first experiment we used collinear sets of dots as targets and we tested the effect of the density and length of the target on its detectability. From Uttal's (1975) results it is known that the detectability of a target increases when either the density or length of the target increases. These results seem to be intuitively obvious and we did not expect to measure different effects. The purpose of this experiment was to describe these relationships in a more systematic way and to extend Uttal's results to different stimuli. Specifically, we used a longer exposure duration and a larger visual angle of the stimuli and we did not mix many target shapes in one session (this last factor is related to familiarity with the target, and it will be directly tested in Experiment 3).

Methods

Subjects. Two of the authors (MSG and ZP) served as subjects. They had prior experience as subjects in psychophysical experiments and had extensive practice for this experiment. MSG was an emmetrope. ZP was a myope and used his normal correcting glasses.

Stimuli. The stimulus consisted of a target and background. The target was a set of collinear, equally spaced dots. The background was a set of similar, randomly distributed dots (see Fig. 4 for an example of a stimulus). The stimuli were displayed on the monitor (1152×900 pixels) of a Sparc 2GS computer. The dots were white (luminance 56.7 cd/m^2). The size of each dot was equal to the size of a pixel. The black background had a luminance of <0.001 cd/m^2. The stimulus occupied a square array of 840×840 pixels. The viewing distance was 150 cm. The stimulus size was 9.4 deg.

The number n of dots in a target of length l and density d was calculated as: $n = l \cdot d + 1$. The length here is specified in pixels, and density is the reciprocal of the distance between neighboring dots (to avoid ambiguity of how the pixels are counted for different orientations of the line, we assume that a pixel has a square shape and the pixel's size is measured as the length of the side of this square). The position and orientation of the target was random from trial to trial. The coordinates of background dots were randomly generated from a uniform distribution. The total number of dots (target and background) in a stimulus was 400.

Procedure. The experiment consisted of two blocks; each block had six sessions. The first block tested the effect of target length for a fixed value of target density (one density level per session). This block is called the *fixed density* condition. The second block tested the effect of target density for a fixed value of target length (one length level per session). This block is called the *fixed length* condition. The order of the blocks was different for the two subjects, and the order of sessions within a block was randomized.

The method of constant stimuli was used. In a given session, for a fixed level of the density (or length), there were seven equally spaced levels of length (or density). The range of levels of density and length were chosen separately for each subject in a preliminary experiment, so that the extremes of the ranges gave rise to close to chance and close to perfect performance. There were 100 trials per level of length (or density). There were also an additional 100 *catch* trials in each session, in which no target was presented. Thus, the total number of trials in a session was 800. Before the experimental session 80 practice trials were used (10 trials per level plus 10 catch trials). The subject was instructed to adopt a response criterion so as to keep the error rate on catch trials as low as possible, but above zero. After each response the subject was given feedback about the accuracy of the response. The subject responded by using the buttons of a computer mouse. The time for response was unlimited. The stimuli in the practice and experimental trials were presented in a random order. After the block of practice trials the subject was informed about the proportion of errors on the catch trials. The subject had an option to repeat the block of practice trials, but this option was rarely used.

Before each trial a fixation cross was presented in the center of the screen. The subject started the trial by pressing a button of a computer mouse. Then, the cross disappeared and after 100 msec the stimulus was presented. The exposure duration of the stimulus was 100 msec. After this time the screen went black and the fixation cross was displayed. The session lasted about 30 min.

The subject sat in a dark room. The monitor was viewed monocularly with the right eye. The subject's head was supported by a chin-forehead rest. The monitor was adjusted in such a way that the subject's line of sight was orthogonal to the screen.

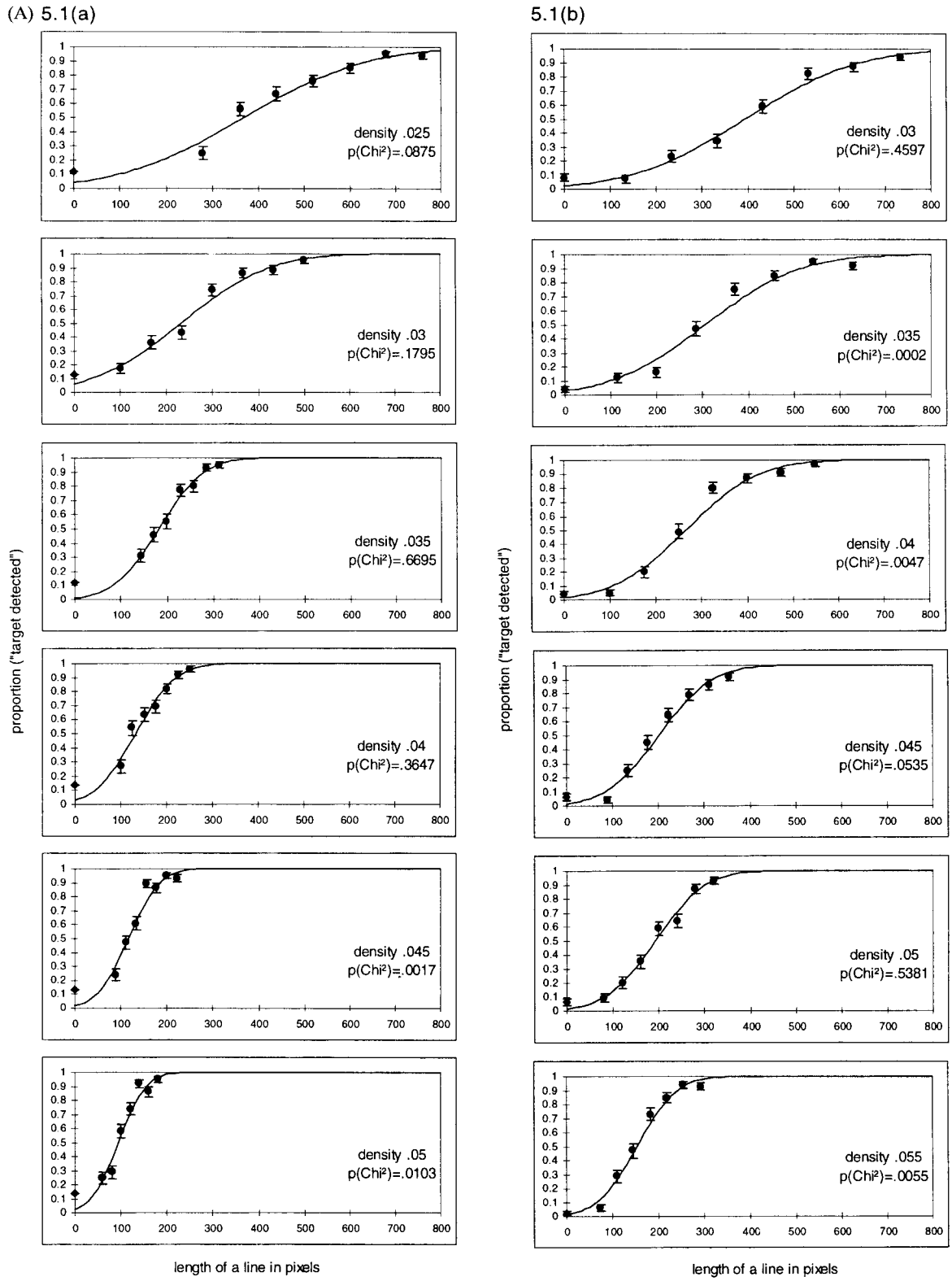

FIGURE 5(A).

Analysis. For each session, the relationship between the proportion of "target detected" responses and the independent variable (density or length) was approximated by a cumulative Gaussian distribution function using Probit Analysis (Finney, 1971). The approximating function was estimated only from points representing trials in which the target was present. The quality of the fit was evaluated using a χ^2 statistic with 5 d.f. (seven data points minus two estimated parameters of the distribution).

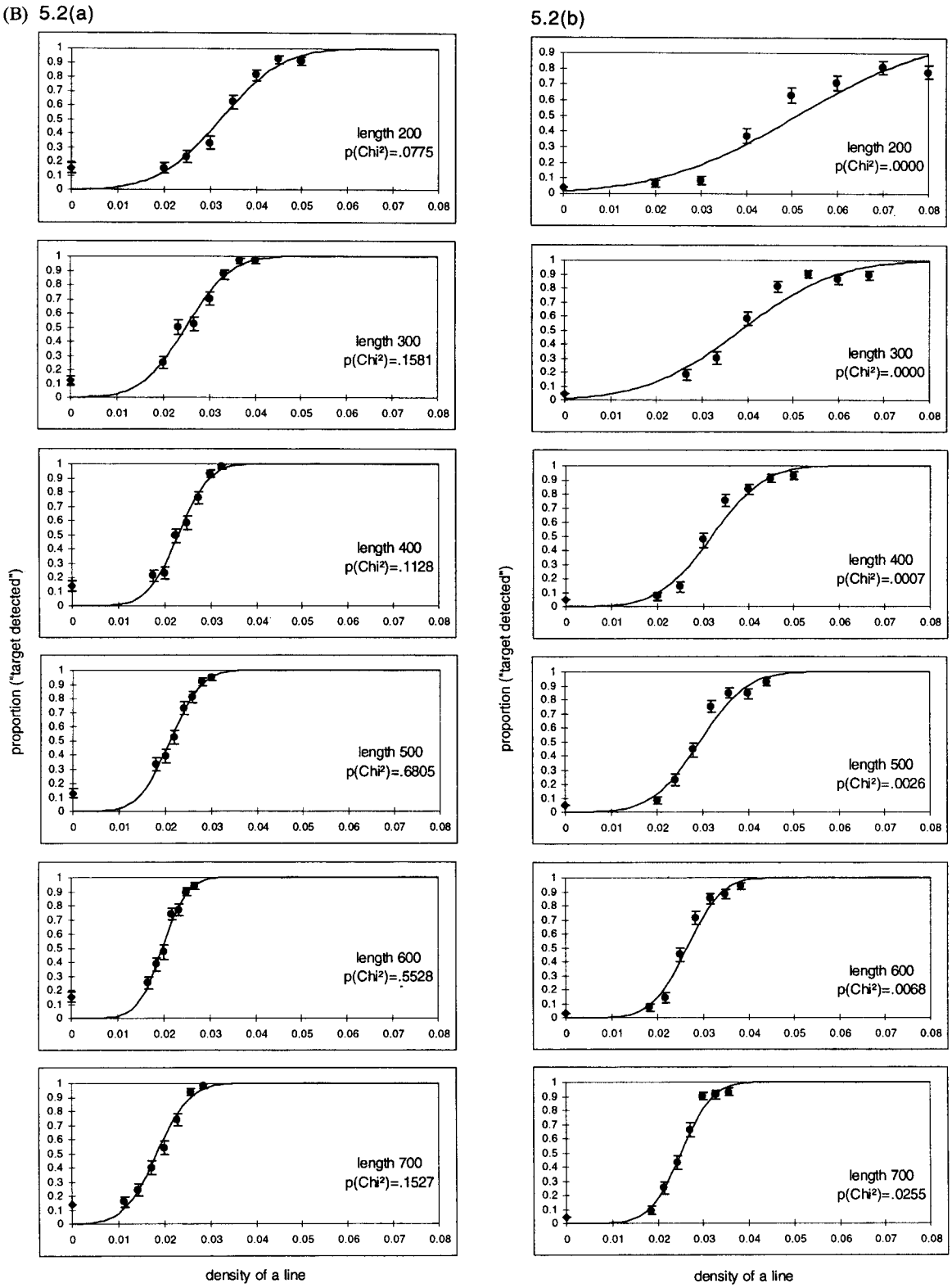

FIGURE 5. (A) Results of Experiment 1 for fixed density condition. The ordinate shows the proportion of responses "target present" and the abscissa shows the length of the target in pixels. Different panels correspond to different levels of the density of dots in the target. Circles represent the trials with target and diamonds represent the trials without target (i.e. catch trials). (a) Subject MSG; (b) subject ZP. (B) Results of Experiment 1 for fixed length condition. The ordinate shows the proportion of responses "target present" and the abscissa shows the density of dots in the target. Different panels correspond to different levels of the length of the target in pixels. Circles represent the trials with target and diamonds represent the trials without target (i.e. catch trials). (a) Subject MSG; (b) subject ZP.

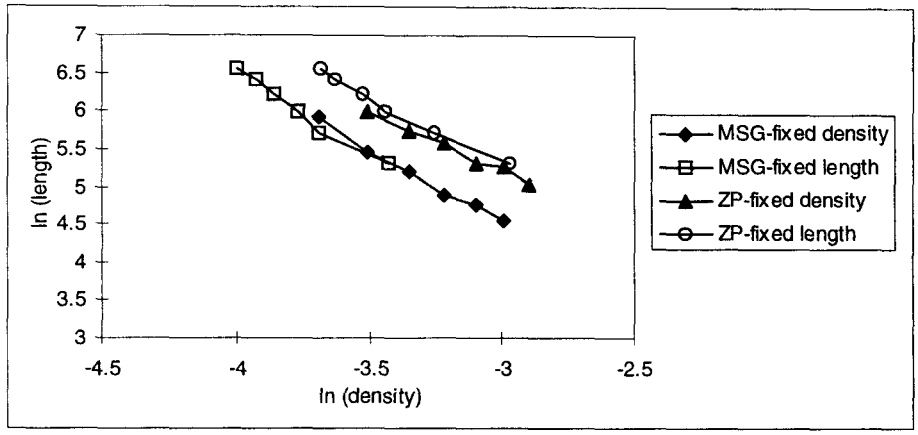

FIGURE 6. Fifty percent detectability curves for Experiment 1. The squares and diamonds represent results from subject MSG, and the circles and triangles from subject ZP. The area above and to the right of the 50% detectability curve represents targets detected in >50% of the trials, whereas the area below and to the left of the curve represents targets detected in <50% of the trials.

Results and discussion

Figure 5(A) shows the results from the fixed density condition, and Fig. 5(B) shows the results from the fixed length condition, for the two subjects. The circles represent trials when the target was present, whereas the diamonds represent catch trials. For each data point there are shown error bars representing the magnitude of sampling error computed as $[(p(1-p)/N]^{1/2}$, where p is the proportion of "target detected" responses and N is the number of observations per data point ($N = 100$). The continuous line is the best fitting cumulative Gaussian distribution function. The probability of obtaining a χ^2 statistic equal to the value computed from the data points or larger is given in each panel. If the fit of the approximating function to the data points is good, the value of the test statistic χ^2 is small, and the corresponding probability is large. Conventionally, a value of this probability >0.1 is assumed to represent a good fit.

For subject MSG the fit was good ($P > 0.1$) in 8 out of 16 conditions and for subject ZP the fit was good in 2 out of 16 conditions. These results suggest the presence of some heterogeneities in the data that could have been produced either:

1. By variability of the dependent variable larger than the variability assumed from the sampling error; or
2. By using an inadequate approximating function.

To check this latter possibility, we repeated the approximation using a logarithmic transformation of the independent variable. This time the fit was good in 5 out of 16 conditions for MSG and in 7 out of 16 conditions for ZP. These results do not allow rejecting one type of independent variable in favor of the other. We conclude that it is more likely that the heterogeneities observed in the data are produced by increased variability of the subject's responses as compared to the theoretical variability predicted from the sampling error. Such increased variability could have come about, for example, by instability in the subject's criterion for the response "target detected".

It is seen from these results that it is easier to detect a line when the line is longer and the density of dots in the line is greater. These results seem to be intuitively obvious and they are consistent with prior results (e.g. Uttal, 1975). To investigate the joint effect of length and density of a line on its detectability, we plotted 50% detectability curves (Fig. 6). This graph shows the relationship between the length of a line and the density of dots in the line in log–log coordinates from both sessions: fixed density and fixed length. For the fixed density condition, the values on the abscissa are the values of density used in the experiment, and the values on the ordinate are the lengths of the lines which were detected in 50% of the trials (i.e. they represent the 50th percentile of the fitted cumulative Gaussian distribution function). For the fixed length condition, the values on the ordinate are the lengths of the targets used in this condition, and the values on the abscissa are the densities of the lines detected in 50% of the trials. Thus, the points plotted in this graph represent 50% detectability curves for each subject across the two conditions (the squares and diamonds for MSG and the circles and triangles for ZP). The area above and to the right of the 50% detectability curve represents targets detected in >50% of the trials, whereas the area below and to the left of the curve represents targets detected in <50% of the trials.

It is seen from this graph that the results from each subject are quite consistent across the two conditions used, namely, the data points corresponding to the fixed density condition and the data points corresponding to the fixed length condition can all be approximated by a single line. Next note that the 50% detectability curves for the two subjects are nearly parallel to one another [their slopes, −1.98 for MSG and −1.79 for ZP, are not significantly different ($P > 0.1$)]. These results suggest that the common perceptual mechanism is represented by the similar slopes of the 50% detectability curves, and that the difference in sensitivity between the two subjects

(the sensitivity of MSG was higher than that of ZP—the 50% detectability curve for ZP is translated to the top and right from that for MSG) is represented by the different intercepts.

In this experiment two features of the target were manipulated, namely, density and length, but the background was kept the same. To evaluate the effect of the background on the subject's performance a follow-up experiment was performed in which both the length and density of the target were kept constant (150 and 0.04, respectively), but the density of the background varied, by using different numbers of dots ranging from 100 to 400. The method of constant stimuli was used with seven levels of the number of background dots. During the "catch trials", when no target was presented, the number of background dots was chosen randomly from the seven levels used in the other trials. ZP was tested in one session (800 trials). It was found that the proportion of "target present" responses was greater when the number of background dots was smaller (the proportion of errors on the "catch trials" was 0.13). The relationship between the proportion of "target present" responses and the number of background dots was approximated by the best-fitting cumulative Gaussian distribution function. The fit, evaluated by the χ^2 test, was quite good ($P > 0.29$). If the logarithm of the number of background dots was used as the independent variable, the fit improved ($P > 0.68$).*

To summarize the qualitative results of this experiment, detectability of a target (a collinear set of dots) is higher when the target is longer, and the density of dots in the target relative to the density of dots in the background is greater. In the next section we will describe our new model based on an exponential pyramid architecture and compare simulation results to the psychophysical results.

THE EXPONENTIAL PYRAMID MODEL OF CURVE DETECTION

First, we will introduce our new definition of smoothness. Then, we will describe the new model of curve detection and illustrate the model's performance as a function of the values of its free parameters. Finally, we will show that this model can generate psychometric curves that fit quite well to those from our Experiment 1.

*We want to point out that the fact that using a logarithmic transformation of the independent variable (length, target density, background density) tended to improve the fit of the approximating functions or simplified their relationships, could be regarded as supporting an exponential pyramid model. This is the case because in the exponential pyramid algorithm the independent variables enter in logarithmic transformations (see Pizlo *et al.*, 1995). Note, however, that there are many other, qualitatively different transformations of the independent variables that could give rise to equally good fits, especially if the range of variation of the variables is relatively small (as is the case in most experiments where detection or discrimination thresholds are measured). Therefore, we will base our evaluation of the model on more direct tests that involve comparison of psychophysical and simulation results.

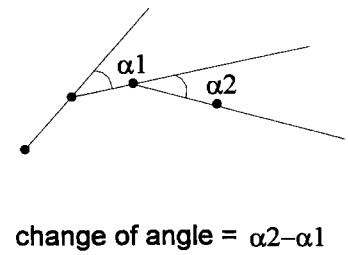

FIGURE 7. Local change of angle ($\alpha_2-\alpha_1$) is defined as the difference between two successive changes of orientation, α_2 and α_1.

A new definition of smoothness

In most prior theories of curve detection, the curve was easier to detect if it was closer to a straight line (see the review in the Prior Research section). This theoretical claim is rooted in the simplicity principle of Gestalt and it is consistent with the results of some psychophysical tests (e.g. Uttal, 1975; Field *et al.*, 1993). The departure from "straightness" was usually defined as the magnitude of curvature (e.g. Sha'ashua & Ullman, 1988). We pointed out, however, that curvature itself is psychologically not plausible because curvature is not scale invariant. The effect of scale on curvature implies that if the observer is approaching a given object (or moving away from it), all curvatures on the retina will change because of changes of retinal sizes. Subjectively, however, the percept does not seem to be affected by the change of distance between the object and the observer. Therefore, we believe that the departure from straightness (or more generally, departure from "smoothness") should be measured by angles rather than curvatures, because angles are scale invariant.

We pointed out in the Introduction that there is an alternative approach (different from the simplicity principle) to figure–ground segregation in general, and the law of good continuation in particular, which involves the likelihood principle. According to this approach, the percept favors interpretations that are likely to occur in everyday life. If this principle is applied to curve (or contour) detection, then the percept should be equally sensitive to a wide range of possible shapes of curves, not just to straight line segments or circles, because the contours found in images of everyday life scenes (e.g. contours of images of animals and persons) do not include many straight line segments or even parts of a circle or an ellipse. Despite the fact that the contours of images in everyday life scenes are characterized by different degrees of departure from straightness, they all share one feature, namely they are piece-wise smooth.

We define smoothness here as a small change of angle along the curve. This definition is illustrated in Fig. 7. Assume that a curve is represented by a set of dots, as in our experiments (note that if a curve is continuous and does not have distinctive dots, one can always divide the curve into short segments and use the endpoints of these segments in lieu of dots). The change of orientation between successive tuples of dots is called an angle (e.g. α_1), and the difference between two successive angles ($\alpha_2-\alpha_1$) is the change of angle. We conjecture that a

curve is smooth (i.e. conforms to the law of good continuation) if the change of angle is small along the entire curve. In a straight line all angles (and changes of angle) are zero. Thus, a straight line is smooth according to this definition. In all circular arcs, the change of angle is also zero everywhere (even though the angle is no longer zero), and, therefore, all circles are smooth, too. In the two lines shown in Fig. 2, the angles are not zero and the changes of angle are not zero either, but the change of angle is small and the curves are perceived as smooth. Before this phenomenological observation is tested (Experiment 2), we describe the details of our model and show that this model can account for the results of our Experiment 1.

The new model

We propose a hierarchical pyramid as a model of curve detection. The structure of the pyramid and the basic operations performed by each node were chosen on the basis of what is known about the anatomy and physiology of the human visual system, as well as on the basis of current knowledge of the perceptual mechanisms of early vision. First, we describe the main properties of the structure of our model.

We used a pyramid with four layers. The bottom layer of the pyramid contained 64 nodes organized in an 8×8 square. The number of nodes at the bottom layer was smaller than the number of pixels in the input image, so each node in this layer received an input from a portion of the image. The ratio of the number of nodes between two successive layers (reduction ratio) was 4. Thus, each "parent node" received input from four "children nodes". In the bottom layer the receptive fields were enlarged by $\pm$ a half of the width (height) of the individual receptive field. As a result, the receptive fields in all layers overlapped by this amount. This allowed avoiding problems in cases when the target to be detected fell on the border of two nodes. Finally, we assumed one global coordinate system in the pyramid and no interactions among nodes in a given layer.

Our pyramid model, whose structure is described above, represents only one particular instance of a more general class of pyramid algorithms. In the general case the reduction ratio can be arbitrary, and it does not even have to be a whole number (Burt, 1981). Receptive fields may or may not overlap. The shape of the receptive field does not have to be a square. In fact, the receptive field sizes and shapes may change as a function of a stimulus (Meer, 1989). Finally, nodes in a given layer may interact and each node may use its own local coordinate system.

Next, we describe rules according to which our pyramid model operates. Each node in the pyramid stores coordinates of dots that are located within its receptive field, the description of *figures* (i.e. sets of dots forming possible targets) detected in its receptive field, along with a measure of "goodness" of each figure. The goodness depends on the figure's length and density. If a node detects at least one figure in its receptive field, this figure, but not the individual dots, is passed up to the next layer for further processing.

A node in the first layer starts processing by analyzing dots in its receptive field to determine if they can form a figure. Our definition of smoothness, which involves the concept of a change of angle, implies that the simplest figure consists of four dots (a quadruple of dots is the smallest set for which a change of angle can be computed). The examination of dots in the receptive fields involves three stages. It begins if the number of dots in the receptive fields is greater than some minimum (in our simulations the minimum number of dots was 6). Otherwise, all dots are passed up for further processing. This requirement prevents the pyramid from losing information: if the dots are very sparse, the target might be missed not because the dots do not form a figure, but because there were not enough dots to verify the existence of a figure or its part. The rules involved in the individual stages are given below:

1. In the first stage, pairs π_{ij} of dots are formed, if the dots P_i, P_j are sufficiently close to one another:

$$P_i,\ P_j \text{ form a pair } \pi_{ij} \text{ if } d(P_i, P_j) \leq \delta$$

where $d(\cdot)$ is a Euclidean distance, and δ is a *maximal distance* (reciprocal of *minimal density*). The restriction that $d(P_i, P_j)$ is small represents the fact that only dots that are close to one another are likely to lead to the percept of a figure (proximity rule). At the same time, this restriction limits the computational time by limiting the number of possible pairs. In the visual system this stage can be performed by simple cells.

2. In the second stage, triplets τ_{ijk} of dots are formed from pairs, if the two pairs have one dot in common and the angle α_{ijk} formed by the two pairs is sufficiently small. Before the magnitude of the angle was evaluated, a random number x_α was added to the actual angle. This random number represented in our simulations perceptual noise in judging angles:

$$\pi_{ij},\ \pi_{j'k} \text{ form a triplet } \tau_{ijk} \text{ if } j = j' \text{ and } |\alpha_{ijk} + x_\alpha| \leq \alpha_{\max}$$

where x_α is a random variable subject to a normal distribution $N(0, \sigma_\alpha{}^2)$ (σ_α is called here *angle standard deviation*), and $\alpha_{\max}$ is the maximal angle (*angle criterion*). In the visual system, this stage can be performed by hypercomplex cells (Dobbins *et al.*, 1989).

3. In the third stage, quadruples κ_{ijkl} of dots are formed from triplets. Two triplets can form a quadruple, if they overlap by two consecutive dots, and the change of angle Δ_{ijkl} is sufficiently small. Before the magnitude of the change of angle was evaluated, a random number y_Δ was added to the actual change of angle. This random number represented in our simulations perceptual noise in judging a difference between angles:

$$\tau_{ijk} \text{ and } \tau_{j'k'l} \text{ form a quadruple } \kappa_{ijkl} \text{ if } j = j' \text{ and } k = k' \text{ and } \Delta_{ijkl} = |(\alpha_{ijk} - \alpha_{j'k'l}) + y_\Delta| \leq \Delta_{\max}$$

where y_Δ is a random variable subject to a normal

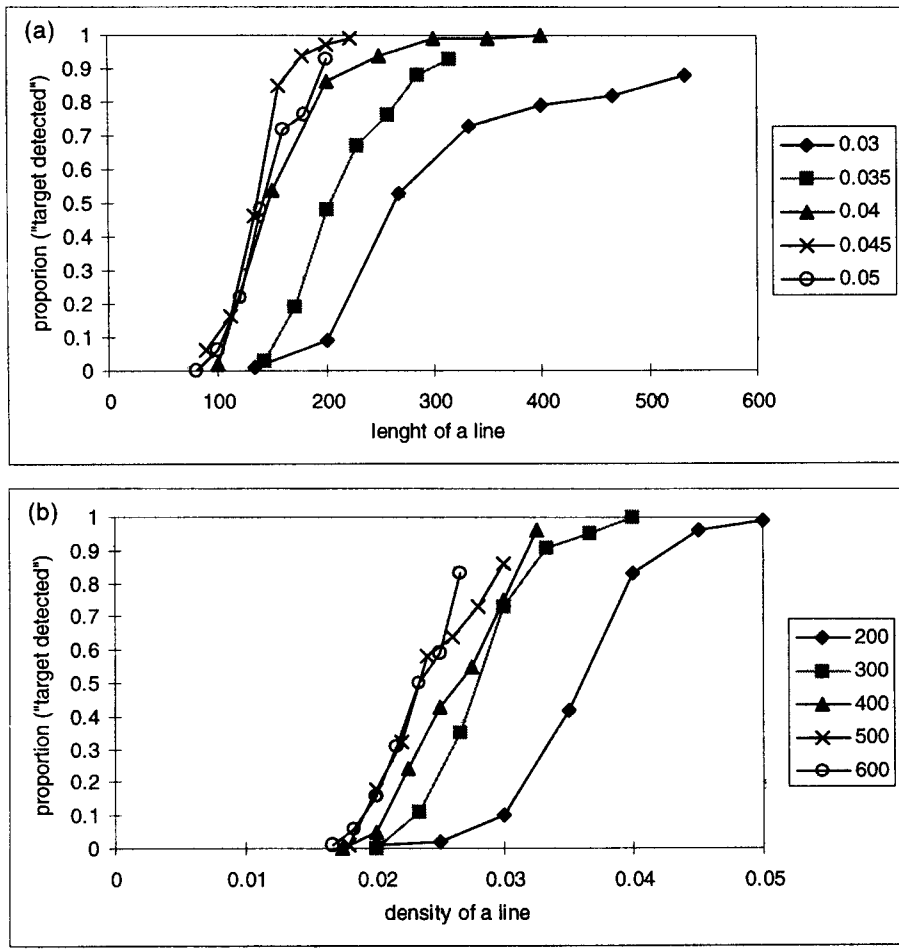

FIGURE 8. "Psychometric" functions generated by the exponential pyramid model using the same conditions as in our psychophysical Experiment 1. The axes in this figure have the same units as in Fig. 5(A) and (B). Specifically, the ordinate shows the proportion of the model's responses "target present". (a) Fixed density condition. The abscissa shows the length of the target in pixels. Different curves represent different levels of the density of dots in the target. (b) Fixed length condition. The abscissa shows the density of dots in the target. Different curves represent different levels of the length of the target in pixels.

distribution $N(0, \sigma_\Delta^2)$ (σ_Δ is called here *change-of-angle standard deviation*), and $\Delta_{\max}$ is the maximal change of angle (*change-of-angle criterion*).

After figures are identified, they are passed up for further processing according to the following rule.

4. If a node on the second (or higher level) receives from its children descriptions of figures $\phi_{i_1,\ldots,i_n}$, it checks whether it is possible to merge pairs of the figures. Such a merge is possible if two figures overlap by more than two dots (starting from one of the endpoints) or they overlap by two dots and the change of angle at the site of the merge is less than the change-of-angle criterion:

$$\phi_{i_1,\ldots,i_n} \text{ and } \phi_{j_1,\ldots,j_m} \text{can be merged if :}$$

$$(a)\ i_{n-k} = j_1, \ldots, i_n = j_{k+1} \text{ and } k \geq 2$$

or

$$(b)\ i_{n-1} = j_1 \text{ and } i_n = j_2 \text{ and } \Delta_{i_{n-2}i_{n-1}i_nj_3} = |(\alpha_{i_{n-2}i_{n-1}i_n} - \alpha_{j_1j_2j_3}) + y_\Delta| \leq \Delta_{\max}$$

Each node orders the figures by using a *goodness* measure and stores the 10 "best" figures for further processing (storing only a small set of figures is consistent with the assumption that the nodes in the pyramid have limited computational and memory capacity). A figure is "good" if it is long and dense. Since we found that the iso-sensitivity curve in Experiment 1 was a straight line in log(length) vs log(density) coordinates, the goodness was measured by the ratio log(length)/log(density). The longer and denser the target, the smaller the value of this criterion (but the larger in absolute value)—see Fig. 6, in which the better figures correspond to points in the upper right corner of the graph. At the apex of the pyramid the goodness of the best figure was compared to a goodness criterion to decide whether the best figure is to be classified as a target.

In the next section we illustrate the performance of the model by analyzing the effects of the length and density of a target on the proportion of correct detections, as well as by analyzing the effects of several parameters of the model on its performance.

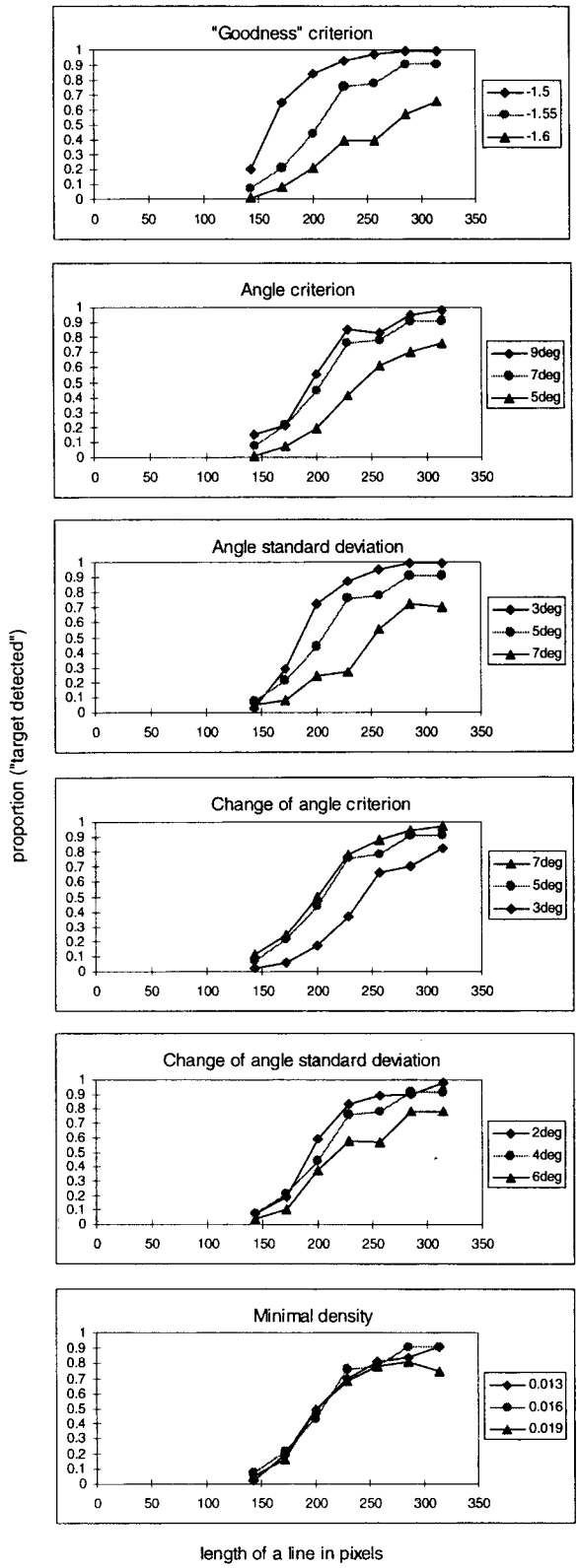

FIGURE 9. Illustration of the effect of model parameters on the model performance. The ordinate shows the proportion of the model's responses "target present" and the abscissa shows the length of the target in pixels. The middle curve (represented by circles) is the same in all six graphs. The other two curves in each panel represent the effect of one parameter on the simulated psychometric function.

Performance of the model

Before we fit the model to the data (next section) we analyze the model's performance in order to examine the family of possible simulation results. If, by changing the model's parameters, one could obtain an arbitrary curve representing the relationship between the proportion of "target present" responses and the independent variable (length or density), then the predictive value of the model would be relatively weak (any result can be accounted for). If, on the other hand, the model can generate only a restricted family of curves, its value as a theory would be much greater (because it can be invalidated).

In the first set of simulations, all parameters of the model were kept constant, and the proportion of "target present" responses as a function of density and length was computed. In these simulations the stimuli were the same as those used in Experiment 1 for subject MSG. Each data point corresponds to 100 trials (as in Experiment 1). The target was always a straight line. The values of the model parameters were as follows: angle criterion (α_{max}): 7 deg, angle standard deviation (σ_α): 5 deg, change-of-angle criterion (Δ_{max}): 5 deg, change-of-angle standard deviation (σ_Δ): 4 deg, maximal distance (δ) between dots that can form a pair: 60 pixels (minimum density: 0.016), goodness criterion: -1.55. These values were chosen in such a way that the model would give rise to psychometric functions in a range of the independent variable (length, density) similar to that produced by the subjects.

Figure 8(a) shows a family of simulated psychometric functions for a fixed density condition and Fig. 8(b) shows the functions for a fixed length condition. It is seen that these simulation curves are qualitatively similar to those obtained from the subjects in Experiment 1 (Fig. 5). This means that our pyramid model is a possible theory of curve detection in the noisy image.

Next, we analyze the effect of the model's parameters on the simulated psychometric functions. Figure 9 shows six panels, each panel having three functions corresponding to three levels of a given parameter. The values of the other parameters were kept constant at their middle values. The targets used in this simulation were straight lines with a fixed density of 0.035. It is seen that all the functions are of the same type, namely, all are approximately monotonic. Changing the parameters resulted in changing the slope or position of the function, or both. Thus, the parameters of the model give rise to systematic changes of the psychometric function. This means that our model cannot reproduce any function. Instead, the model always produces a function having a shape similar to the shape of a cumulative Gaussian distribution function. The next section will test how closely the simulated functions can reproduce (approximate) the psychometric functions measured in Experiment 1.

Fitting the model to the data

Fitting the model to the data consisted of determining the optimal values of the parameters that led to the best fit

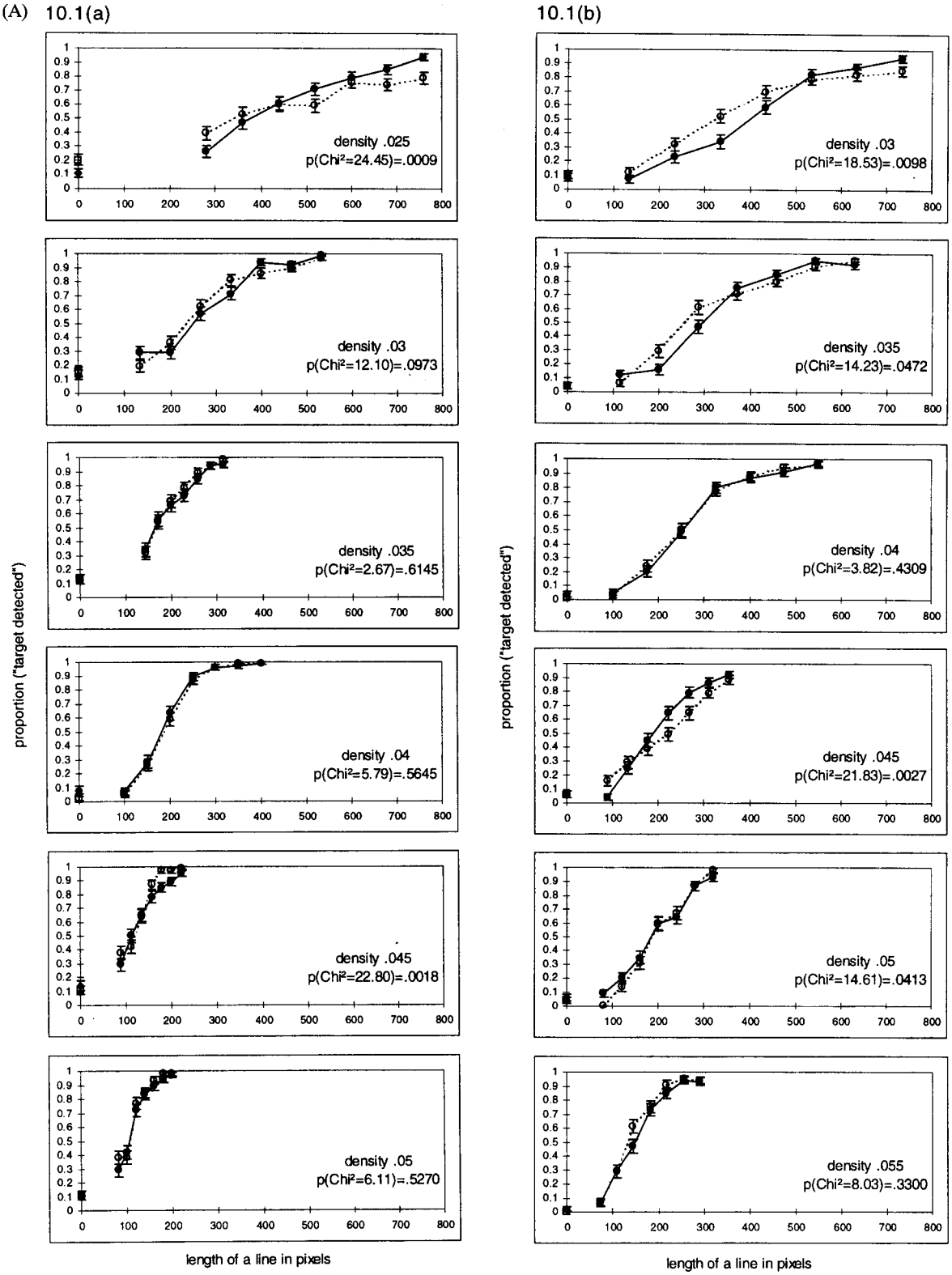

FIGURE 10(A).

(as measured by a χ^2 test) of the simulated psychometric functions to those obtained in Experiment 1. Since in Experiment 1 only straight lines were used as targets, this optimization did not involve the change-of-angle criterion (Δ_{max}) or the change-of-angle standard deviation (σ_Δ) (these two parameters will be used in the simulations of Experiment 2). As a result, only the following parameters were subject to optimization:

1. Maximal distance (δ) between two dots that could form a pair in a figure (reciprocal of minimum density);

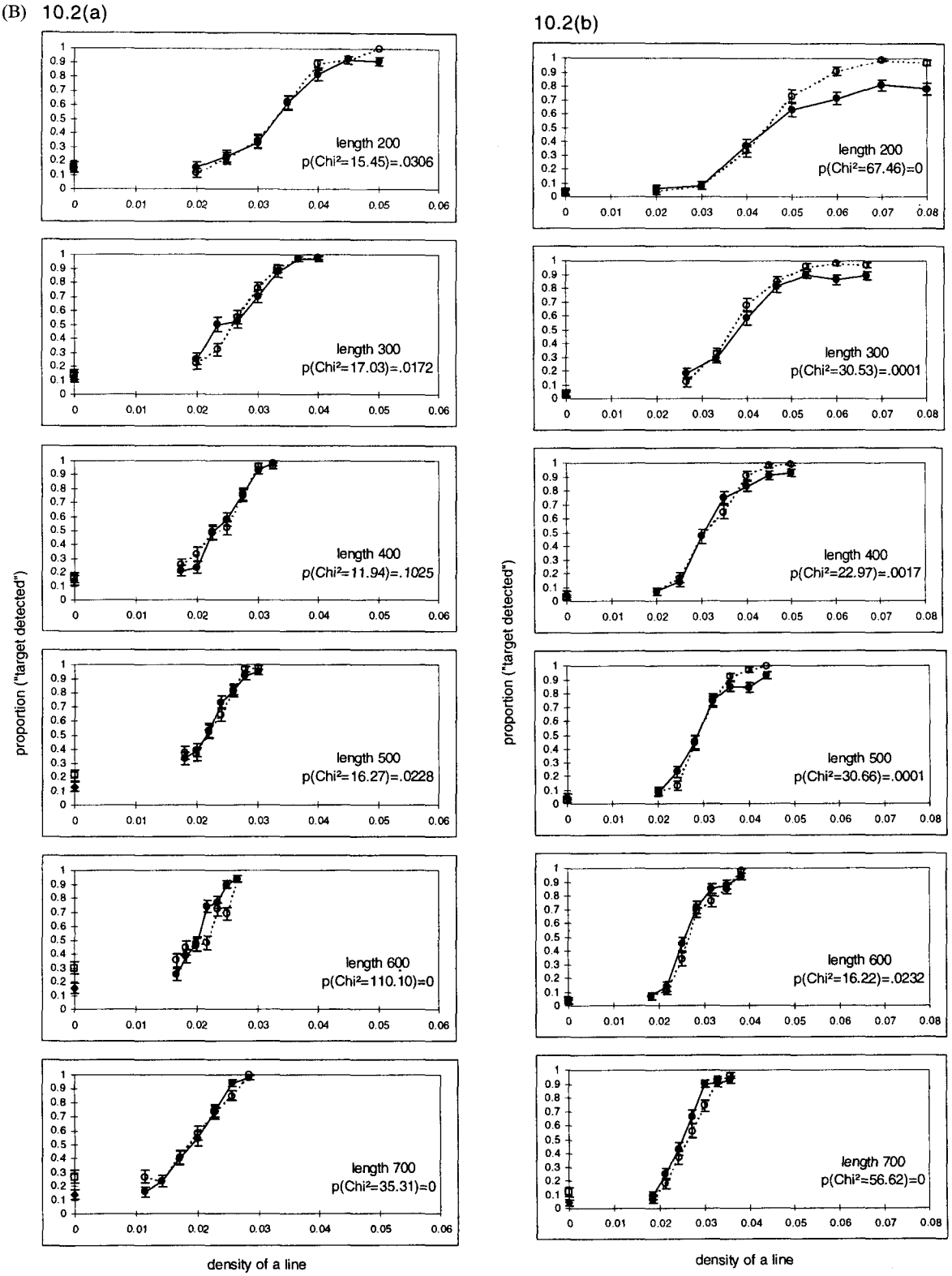

FIGURE 10. (A) Simulated and psychophysical psychometric functions for the fixed density condition from Experiment 1. Filled circles (connected by a solid line) and a diamond represent psychophysical results [in Fig. 5(A)] and open circles (connected by a dashed line) and a square represent simulated results. (a) Subject MSG; (b) subject ZP. (B) Simulated and psychophysical psychometric functions for the fixed length condition from Experiment 1. Filled circles (connected by a solid line) and a diamond represent psychophysical results [shown in Fig. 5(B)] and open circles (connected by a dashed line) and a square represent simulated results. (a) Subject MSG; (b) subject ZP.

2. Angle criterion (α_{max});
3. Angle standard deviation (σ_α);
4. Goodness criterion.

The optimization was performed separately for each subject because it was reasonable to assume that the optimal values of the parameters in the model may be different for different subjects. Among the four parameters listed above, the first three are likely to characterize the structure and the functioning of the visual system itself, whereas the fourth parameter (the goodness criterion) is likely to correspond to the subject's response criterion. Therefore, the values of the three parameters were estimated only once (for one psychometric function) and then were kept constant. The value of the fourth parameter (the goodness criterion), on the other hand, was optimized for each psychometric function.

Thus, the optimization was performed in two stages. In the first stage, optimal values of all four parameters were estimated by fitting the simulated curve to the psychophysical one for the middle (third) session of the fixed density condition. The optimal values of the three parameters (these values were kept constant in the remaining optimizations) were as follows (for MSG and ZP, respectively):

1. Minimum density ($1/\delta$): 0.013 and 0.012;
2. Angle criterion (α_{max}): 6 and 4 deg;
3. Angle standard deviation (σ_α): 9 and 10 deg.

The remaining 11 psychometric functions (five from the fixed density condition and all six from the fixed length condition) from the first experiment were fitted by using only one free parameter, namely, the goodness criterion. The results are shown in Fig. 10. The data points representing the psychophysical experiment are represented by filled symbols: circles joined with a solid line for trials with target, and diamonds for catch trials. The data points representing the simulation experiment are represented by open symbols: circles joined with a dotted line for trials with target, and squares for catch trials. It is seen that the fit was very good for the third session in the fixed density condition, for which four parameters were subject to optimization (in this case the number of degrees of freedom was four). The fit was also good for several other sessions from the fixed density condition (in these cases the number of degrees of freedom was seven). Note that the quality of the fit is similar to that obtained by fitting the cumulative Gaussian distribution function to the psychophysical psychometric functions (see Fig. 5). In the fixed length condition, where the simulated functions were produced by using estimates of three parameters from the fixed density condition, the fit, as measured by a χ^2 test, was worse. It is quite possible that the fit could be improved if more than one parameter was allowed to be optimized. It is seen, nevertheless, that the shapes of the simulated psychometric functions are similar to the shapes of the psychophysical psychometric functions. We conclude that the simulation results produced by our new model are consistent with the psychophysical results. It is quite possible, however, that similarly good fits could have been obtained by other models as well [e.g. Uttal's (1975)]. Therefore, new experiments are needed that allow the different models to be told apart. Such experiments will be presented in the next two sections.

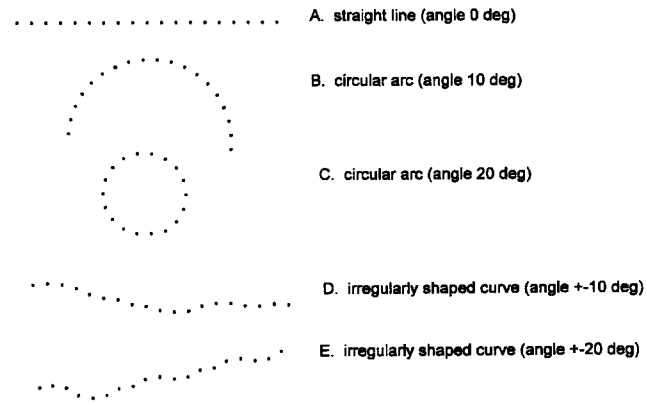

FIGURE 11. Targets used in Experiment 2. The targets have different shapes, but they all have the same length and density.

EXPERIMENT 2: DETECTION OF CURVED LINES

Existing definitions of smoothness (or good continuation) of a curve can be classified by two parameters: local angle and local change of angle. According to most prior theories, a curve is smooth if the local angles are small along the curve (e.g. Uttal, 1975; Grossberg & Mingolla, 1985a,b). As a result, in these theories, a straight line is the easiest to detect and detectability should deteriorate if the local angle increases (e.g. a small circle should be more difficult to detect than a large one). According to other theories (e.g. Smits & Vos, 1986; Yuen *et al.*, 1990), a curve is smooth if the local angle is constant along the curve (i.e. the change of local angle is zero). In these theories all circles and straight lines should be equally easy to detect and detectability should deteriorate if the shape of the curve is irregular. On the contrary, according to our new definition of smoothness, detectability should be equally good for a wide range of shapes: straight lines, circles, and irregularly shaped curves, provided the local change of angle is small. To test which of the theories, if any, is psychologically plausible, we used the following shapes of targets (see Fig. 11):

1. Straight line;
2. Circular arc with local angle 10 deg;
3. Circular arc with local angle 20 deg;
4. Irregularly shaped curve with maximal change of local angle 20 deg;
5. Irregularly shaped curve with maximal change of local angle 40 deg.

All targets had the same density of dots and the same length. Therefore, any differences in the subjects' performance can be attributed to the shape itself, rather than to other unrelated factors.

Methods

Subjects. Three subjects, including two of the authors, were tested. Subjects MSG and ZP served in the first experiment. The third subject, BJ, was naïve about the hypothesis being tested. BJ was a slight myope and he did not usually wear his corrective glasses. Therefore, he was tested in the experiment without glasses. He did not have previous experience as a subject in psychophysical experiments. He was given extensive practice for the current experiment.

Stimuli. The length of the target and the density of dots in the target were constant for all sessions for a given subject. Subjects ZP and BJ were presented with targets of length 400 and density 0.045. Subject MSG was presented with targets of length 600 and density 0.03, which were more difficult than the targets used by the other two subjects. Testing MSG with more difficult stimuli led to similar performance in all subjects, due to the fact that MSG had greater detectability (see the Results section of Experiment 1 for comparison of the detectability of MSG and ZP). The number of dots in each target was 19 (in all sessions and for all three subjects), except in the case of a circular arc with local angle 20 deg [Fig. 11(C)]. In this case the circle was closed and, as a result, the first and the last dot coincided.

Circular arc targets [Fig. 11(B and C)] were generated by randomly choosing the starting point and the orientation of the segment represented by the first two dots. The remaining dots were generated on the circumference of the circle in such a way that the angle formed by three consecutive dots was constant and equal to the given angle α (we used α equal to 10 deg and 20 deg). The shape of the irregularly shaped targets [Fig. 11(D and E)] was random from trial to trial and was obtained by using the following method. The first two dots in the case of an irregularly shaped target were generated identically to those in the case of a circular target. The remaining dots were generated in such a way that the angle formed by three consecutive dots was equal to plus or minus the given angle α (again, we used α equal to 10 and 20 deg). As a result, the local change of angle could take values 0, $+2\alpha$, or -2α.

Procedure. The experiment consisted of five sessions. The order of the sessions was randomized for each subject. In each session only one target was used. The shape of an irregularly shaped target Fig. 11(D and E) was random and changed from trial to trial.

The method of signal detection with confidence rating was used. In a given session there were 300 trials in which a target was present, and 300 trials without a target. Before each session 60 practice trials were presented to the subject; in half of the trials the target was present.

The subject's task was to detect a target. After the stimulus disappeared the subject was presented with five possible responses: "noise", "probably noise", "uncertain", "probably signal", and "signal".

Signal detection theory for the data containing confidence ratings was used to analyze the results of the experiment (Macmillan & Creelman, 1991). The analysis provided the detectability measure d' and its estimated standard deviation.

Results

The results are shown in Fig. 12. Each panel shows the results of all sessions for one subject. On the abscissa the type of target used in a given session is indicated (straight line, circular arc with 10 deg local angle, circular arc with 20 deg local angle, irregularly shaped curve with ± 10 deg local angle, irregularly shaped curve with ± 20 deg local angle). For each session (target) the value of d' is represented by the height of a bar. Additionally, error bars, with heights equal to $\pm$ 1 SD of d', are marked.

It is seen that all subjects produced the same pattern of results. The straight line was easiest to detect.* The next three targets: circular arc with local angle 10 deg, circular arc with local angle 20 deg, and irregularly shaped curve with local angle 10 deg, show similar detectability for each subject. The differences among these three conditions are small and comparable to the standard deviations of d'. Finally, the session with irregularly shaped curves with local angle 20 deg produced the poorest performance.

Discussion

We will now compare these results to the predictions of prior theories of the law of good continuation. First, the fact that the detectability of an irregularly shaped curve

*The detectability d' for a straight line, estimated in this experiment by the use of signal detection theory, agrees quite well with the discriminability estimated in Experiment 1, where the method of constant stimuli was used. Consider first the results of MSG in Experiment 1 in the fixed length condition for length 600 (i.e. the length used in the present experiment). The proportion of "target present" responses on catch trials was 0.15. This corresponds to the standard normal variable $z = -1.04$. Next, consider a target with density 0.03 (the density used in the present experiment). A target with this density was not used in the session with length 600; therefore there is no measurement of performance for such a target. However, the performance (i.e. the proportion of "target present" responses) can be estimated from the best fitting cumulative Gaussian distribution function. In this session the fitting curve had mean 0.0197 and standard deviation 0.0043. This means that the density 0.03 corresponds to a z score of +2.40. As a result, the discriminability d' between noise and a target of length 600 and density 0.03 is estimated as 3.44. This estimate agrees quite well with the d' estimated from the present experiment for MSG and the straight line condition (see Fig. 12). Next, we perform a similar comparison for ZP. Consider the fixed length condition with length 400 and take the data points representing catch trials and a target with density 0.045. The proportion of "target present" responses on catch trials was 0.05, which corresponds to $z = -1.64$, and the proportion of "target present" responses for density 0.045 was 0.91, which corresponds to $z = 1.34$. Thus, the d' estimated from Experiment 1 is 2.98. Again, this estimate is quite close to the d' measured in the present experiment (see Fig. 12). This agreement between the results obtained in the two experiments means that the subjects' performance did not depend on the psychophysical method used (constant stimuli in Experiment 1 vs signal detection in Experiment 2).

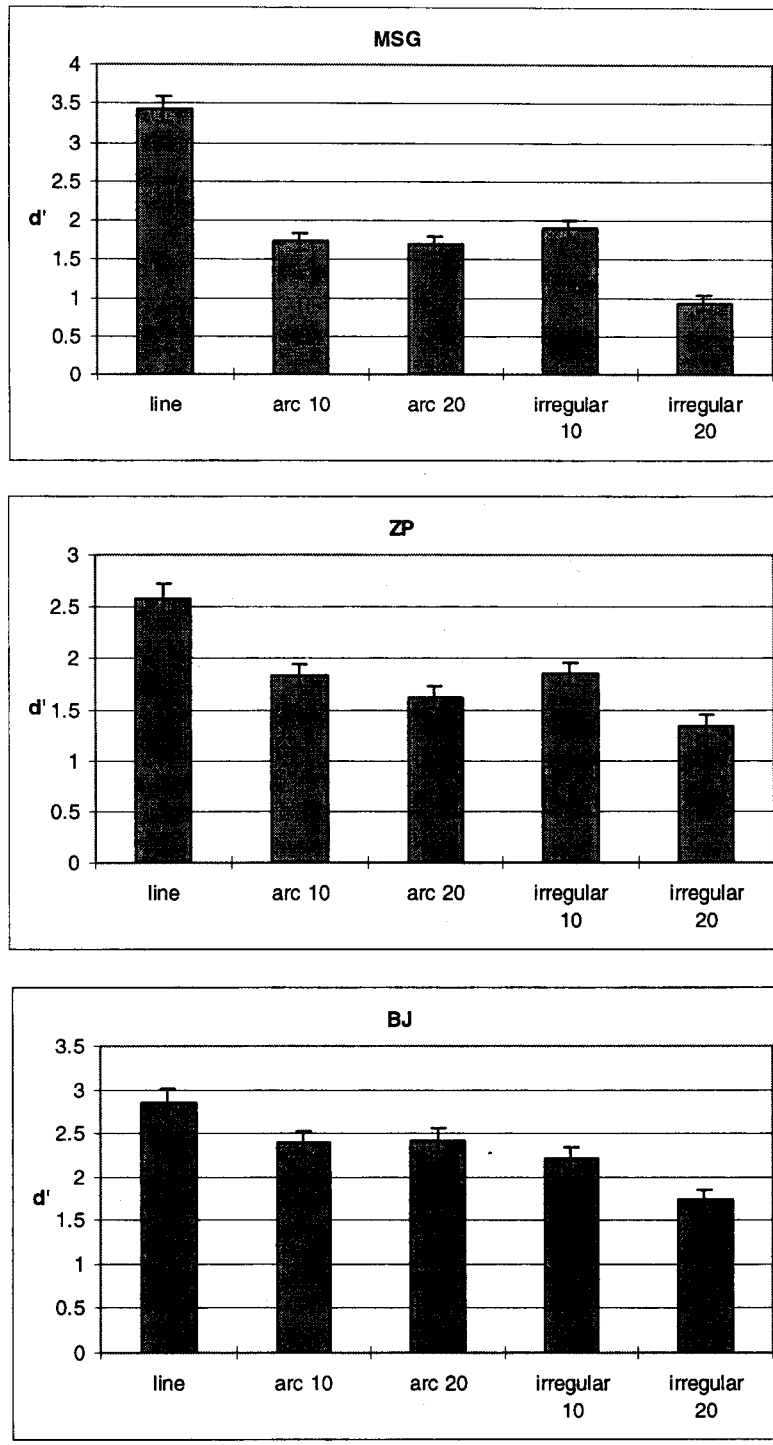

FIGURE 12. Results of Experiment 2. Individual bars represent different targets (shown in Fig. 11). Different panels show results for different subjects.

with a local angle ± 10 deg is not different from the detectability of a circular arc implies that good continuation (or smoothness) is not equivalent to simplicity of the curve representation. An irregularly shaped curve changes its direction randomly; to describe the shape of such a curve the number of parameters needed would be comparable to the number of dots in the curve, whereas the shape of a circular arc can be described by only two parameters: the local angle and the length of the arc. Thus, our results contradict the simplicity principle and all prior theories based on this principle. The only property common to a circular arc and to an irregularly shaped curve with local angle ± 10 deg is that these curves do not change direction rapidly, i.e. the change of local angle is small along the curve (in this experiment this change of local angle was not >20 deg).

When the change of local angle was greater (40 deg), as in the case of an irregularly shaped curve with local angle ± 20 deg, detectability was poorer. These results are consistent with our new definition of smoothness and they seem to reflect the operation of rule 3 (see *The New Model* section).

Next, consider the fact that the performance was the same for the two circular arcs. This result implies that the value of the local angle itself (or the value of curvature) has no effect on the detectability of an arc. Therefore, good continuation (or smoothness) of a curve does not seem to involve local angle (or curvature) and thus seems to reflect in this case the absence of rule 2 (*The New Model* section) or the use of a liberal criterion α_{max} in this rule. This result is consistent with our definition of smoothness and it contradicts those prior theories which claimed that the greater the departure from straightness, the worse is the performance.

Next, consider a straight line, where performance was best. This superiority of the straight line target was not predicted by our definition of smoothness. According to our definition, if the change of local angle is small, the detectability should be high, and the same should be true for different curves. Can our pyramid model account for this higher performance in the case of a straight line? If the model "knows" that the target is a straight line, then it can use the constraint on the bottom layers of the pyramid, that all local angles are close to zero. This constraint is in fact represented by rule 2 (*The New Model* section) (recall that using this rule allowed our model to achieve the same (high) performance as the subjects did in Experiment 1). A question is, however, why the angle criterion produced by knowledge about the local angle characterizing the target can improve performance in the case of straight lines, but not in the case of circular arcs? Consider a circular arc and assume that the local angle α_0 formed by a triplet of consecutive dots is known. If the orientation of the circular arc is unknown, a given pair of dots does not determine the position of the third dot uniquely because the angle α_0 can be positive or negative. In other words, given a pair of dots from a circular arc, it is not known whether the arc was generated clockwise or counterclockwise. In such a case if information about α_0 is to be used, it would correspond to the following modification of rule 2 (rule 2a):

$$|\alpha_{ijk} + x_\alpha \pm \alpha_0| \leq \alpha_{max}$$

This inequality means that the algorithm includes in the analysis all triplets of dots whose angles $(\alpha_{ijk} + x_\alpha)$ are within the region having angular magnitude of $4\alpha_{max}$ ($\alpha_0 \pm \alpha_{max}$ and $-\alpha_0 \pm \alpha_{max}$). In the case of a straight line ($\alpha_0 = 0$) the algorithm includes in the analysis only those triplets of dots whose angles $(\alpha_{ijk} + x_\alpha)$ are within the region having angular magnitude of $2\alpha_{max}$ ($\pm\alpha_{max}$). Thus, the region that has to be analyzed is two times smaller in the case of a straight line as compared to the case of a circular arc. In other words, in the case of a straight line, predictability of the location of the next dot in the line is two-fold greater than that in the case of a circular arc with unknown orientation. This means that the knowledge of local angle in the latter case could be used, but the benefit from such top-down processing may be small enough to be overshadowed by the performance based on bottom-up processing. This reasoning can be tested by measuring performance for a circular arc with fixed orientation. In such a case one can take the direction of the chord of the arc as the reference orientation (this orientation would be constant and known to the subject). This direction would determine the sign of α_0 in rule 2a, and, as a result, the uncertainty (and thus the performance) would be similar or equal to the uncertainty (and performance) in the case of a straight line with unknown orientation. Such a test will be described in the next section.

Finally, consider how top-down effects could account for the apparent contradiction between Uttal's results, where he observed that the greater the departure from straightness, the poorer the performance, and our results, where we failed to observe such an effect (see the comparison between circular arcs with local angle 10 and 20 deg). In Uttal's experiments straight line targets and curved targets were all presented in random order in a single session. In such a case, if the subject used, at least in some trials, a criterion that the local angle is close to zero (rule 2) (as would be the case if only a straight line target was used), performance would be systematically affected by the departure from straightness. This is what Uttal observed. In our experiment, on the other hand, straight line targets were never mixed with curved targets in a single session and, therefore, the subject did not have to use two different criteria at the same time. As a result, when a curved line was the target, the subject was not likely to use the criterion for straight lines, and conversely, when a straight line was the target, the subject was likely to use the criterion for straightness in all trials, which could give rise to high performance.

In the next section, we present the results of simulations in which our pyramid model was tested with the same stimuli as those used in Experiment 2.

Simulations

Since the model has already been tested for the case of straight line targets (see Fig. 10), the present simulations were performed to test the model's capability to account for the subjects' performance in the case of curved lines. Therefore, the angle criterion (α_{max}) and angle standard deviation (σ_α) were not used in the present simulations. Instead, the change-of-angle criterion (Δ_{max}) and the change-of-angle standard deviation (σ_Δ) were used. Figure 13 shows the results of simulations for three conditions represented by different pairs of values for the change-of-angle criterion and change-of-angle standard deviation: 10 and 8 deg, 20 and 10 deg, and 25 and 20 deg, respectively. These three conditions were chosen in such a way that they gave rise to qualitatively different results. In the first case, both the criterion and standard deviation for change of local angle were small. In the second, the standard deviation was small, but the criterion was large. In the last simulation, both the

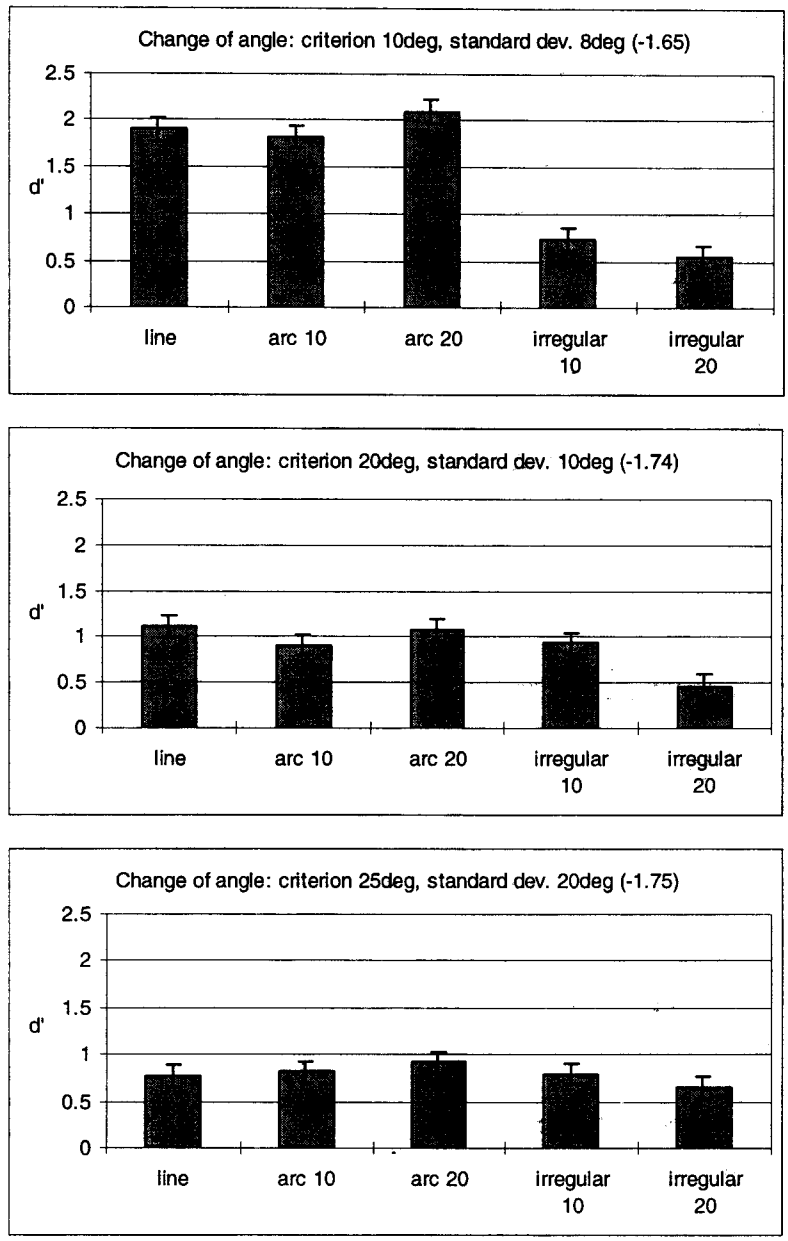

FIGURE 13. Results of simulations for Experiment 2. Individual bars represent different targets. Different panels show different combinations of the model parameters: change-of-angle criterion and change-of-angle standard deviation. The numbers in parentheses are values of the goodness criterion.

criterion and standard deviation were large. The case when the criterion is small and the standard deviation is large is not interesting here because this would lead to a situation where many (or most of) the targets are missed. As a result, detectability would be close to the chance level for all targets.

In the simulations, the target had length 600 and density 0.03 (the same as used by MSG). The session consisted of 300 images with target and 300 images with noise. The proportion of the model's "target present" judgments for trials with and without target were recorded and used to compute the detectability measure d'.

The results of the simulations are shown in Fig. 13. The form of this graph is identical to the form of the graphs shown in Fig. 12. Each panel in Fig. 13 represents one simulation for one pair of values for the criterion and standard deviation of change-of-angle. It is seen that the pattern of results depends on the values of the two parameters and that the results shown in the middle panel most closely resemble the pattern of results obtained by the subjects. Specifically, the model's detectability of an arc with local angle of 10 deg, an arc with local angle of 20 deg and an irregularly shaped curve with local angle ± 10 deg are similar (the differences in d' among these conditions are small and comparable to standard errors).

At the same time, d' for an irregularly shaped curve with local angle ± 20 deg is clearly lower. This pattern of results is identical to that obtained by the subjects. This agreement between the psychophysical and simulation results provides support for our model of curve detection, along with our new definition of smoothness.

There is, however, one psychophysical result which was not reproduced in our simulation. Namely, the subjects were able to detect the straight line target better than any other target (see Fig. 12). This result was not obtained in our simulations because the simulations were based on the new definition of smoothness, and, in this definition, the straight line target is not different from other smooth line targets. This comparison of the psychophysical result with the simulation results suggests that detecting straight lines may involve using an additional constraint on local angle, which leads to a more efficient mechanism (as pointed out in the previous section). Such a constraint could have been used by our subjects because in the session with a straight line target, no other target was used and the subjects knew this fact. As we already showed in the section describing our model, where the psychophysical psychometric functions were approximated quite well by the simulation psychometric functions, this constraint can be implemented in our pyramid model as well, and it leads to performance as good as that of the subjects.

To summarize, we can conclude that our simulation model can account quite well for the results of our psychophysical experiments on detection of straight and curved targets. Furthermore, our model allows for implementing some top-down constraints that may result from knowledge about properties of the target to be detected. This fact agrees with our psychophysical results. First, in Experiment 2, the detectability of a straight line was clearly better then the detectability of other lines. This result shows that some information about the shape of the target can be used to improve detectability. Second, we showed in our Experiment 2 that the detectability of circular arcs (whose shape was constant in a session and known to the subject) was the same as detectability of an irregularly shaped curve with local angle ± 10 deg (whose shape was random and unknown to the subject). This result suggests that knowledge of shape is not always used, and this result was consistent with our new definition of smoothness, according to which an arrangement of dots is classified as a smooth line (and thus detected) if the change of local angle is small along the entire curve. These two groups of results show that curve detectability conforms to the new definition of smoothness, but at the same time it allows for using knowledge about some properties of the target, provided that the knowledge refers to local properties of the target.

To further study the role of knowledge about a target, Experiment 3 was performed. In this next experiment the effect of knowledge of a local property (orientation) and a global property (shape) of a curved target on its detectability were tested. Experiment 2 already revealed the lack of an effect of knowledge about a shape on its detectability. However, this result was obtained by comparing detectability for circular arcs with that for irregularly shaped curves. As a result, the knowledge about the shape was confounded with the target shape itself. To test the effect of knowledge about local and global properties of a target, unconfounded with the target's shape, Experiment 3 tested the effect on detectability of:

1. Knowledge of the orientation of a circular target;
2. Knowledge of the shape of an irregularly shaped curve; and
3. Knowledge of both the shape and orientation of an irregularly shaped curve.

EXPERIMENT 3: THE ROLE OF KNOWLEDGE OF THE SHAPE AND ORIENTATION OF A TARGET

Consider first theoretical predictions of how knowledge about a target can be used. In general, such knowledge can be used in two ways: globally, while making a decision about whether a seen figure is the target, or locally, to restrict the number of accepted fragments of figures. The two possibilities of using knowledge about the target have different implications. If the knowledge is to be used globally, the figure found in the image must be compared with the expected target. In this method, there is no restriction on the type of information that may be useful. It is clear, however, that this method cannot be implemented in the pyramid algorithm because global information about the target is not available to nodes on the layers that are lower than the layer of the *root* node (the root is a node which can "see" the entire stimulus). Clearly, the pyramid algorithm can use only local information. More precisely, only such local information can be used, which can be translated into rules that can be applied locally, so that many nodes in the pyramid structure, including the nodes that do not "see" the entire stimulus, can use this information to restrict the number of possible targets. As a result, the pyramid algorithm predicts benefit from knowing both the shape and orientation of a target and no benefit from knowing only the shape of an irregularly shaped target.

Methods

Subjects. Three subjects, including the two authors who served in the first and second experiments, were tested. The third subject, PG, was an emmetrope. He did not have any prior experience as a subject in psychophysical experiments. He was given extensive practice for the current experiment.

Stimuli. Only two types of targets were used: a circular arc with local angle of 10 deg and an irregularly shaped curve with local angle of ± 20 deg. The targets had constant length and density, the same for all sessions for a given subject. Subjects MSG and PG were presented with targets of length 600 and density 0.03, whereas subject ZP was presented with targets of length 400 and density 0.045.

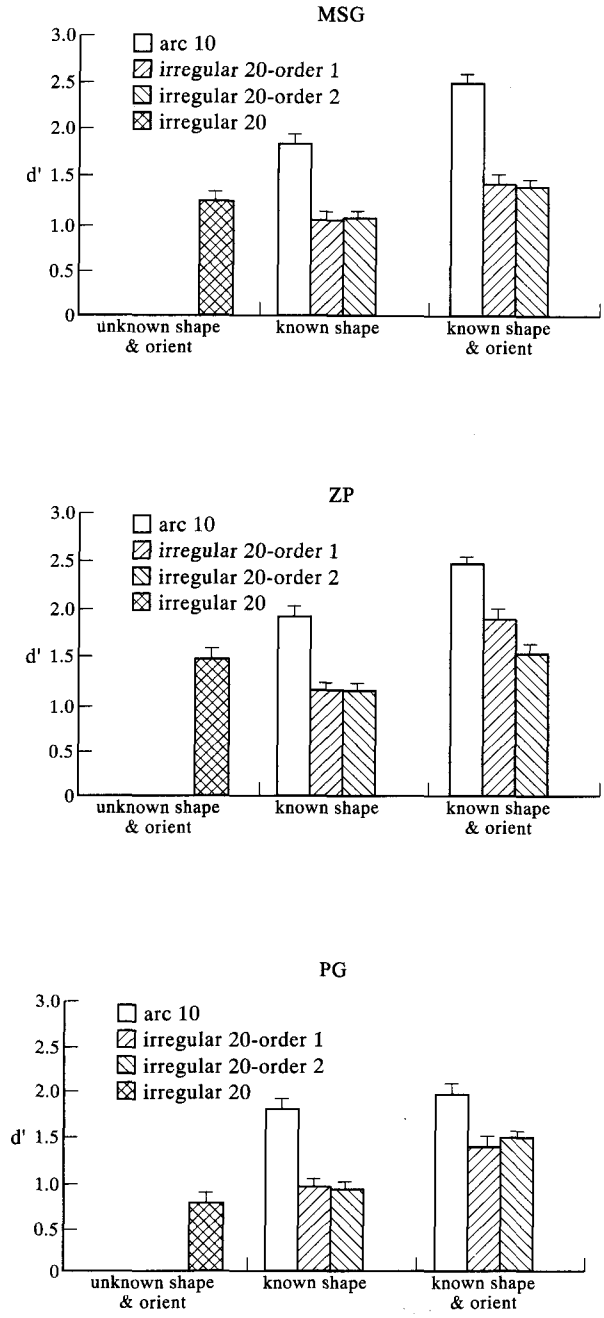

FIGURE 14. Results of Experiment 3. The three groups of bars represent the three levels of familiarity with the target. The individual bars represent the following conditions: (from the left): irregularly shaped curve with unknown shape and orientation, circular arc with unknown orientation, irregularly shaped curve with known shape but unknown orientation (order 1, order 2), circular arc with known orientation, irregularly shaped curve with known shape and orientation (order 1, order 2).

Procedure. The experiment consisted of five sessions. The order of the sessions was randomized for each subject. In each session only one type of target was used. The following targets were used:

1. Circular arc (local angle 10 deg) with randomized orientation (as in Experiment 2);
2. Circular arc (local angle 10 deg) with fixed orientation;
3. Irregularly shaped curve (local angle ± 20 deg) with randomized shape and orientation (as in Experiment 2);
4. Irregularly shaped curve (local angle ± 20 deg) with fixed shape, and randomized orientation;
5. Irregularly shaped curve (local angle ± 20 deg) with fixed shape and orientation.

The position of the target was randomized in all conditions. Note that conditions 1, 2, 4, and 5 use a constant and known shape (the subjects were familiarized with the shape before the session). Condition 3 uses a randomly changing shape of the same type as in condition 4 and 5. Conditions 1, 3, and 4 use random orientations of the target and conditions 2 and 5 use fixed orientations of

the target. To be able to compare conditions 4 and 5 it was important that they used a curve with the same shape. Otherwise, the advantage of condition 5 over 4 could be related to the fact that the shape in condition 5 was coincidentally easier to remember, rather than to the fact that orientation was fixed. Note, however, that using the same shape in conditions 4 and 5 could give rise to an order effect, for example, if condition 5 was run after condition 4, the better detectability in condition 5 could be attributed to the fact that the subject had more experience with this curve while running condition 5. To unconfound the effect of order, shape and orientation, conditions 4 and 5 were repeated twice for each subject: in order 4, 5 and in order 5, 4. In each such set of two sessions the same shape was used.

At the beginning of each session the subject was shown a sample target for unlimited inspection. The subjects were informed what properties of the target would stay constant in a given session, and were instructed to carefully examine and remember the sample target. The rest of the details of the experimental design were the same as in the second experiment.

Results and discussion

The results are shown in Fig. 14. We will concentrate here on describing and discussing two results that were observed for all three subjects. First is the lack of effect of knowing the shape of an irregularly shaped target on its detectability when its orientation is unknown (condition 3 vs condition 4). Second is the improvement of detectability in the case of a target whose shape and orientation are known (conditions 2 and 5), as compared to a target whose shape is known but whose orientation is unknown (conditions 1 and 4, respectively).

Consider now in more detail the case of a circular arc (conditions 1 and 2). Condition 1 was identical to one of the conditions in Experiment 2. As expected, the performance in this condition of the two subjects who participated in all three experiments (MSG and ZP) was identical in both experiments. Next, consider condition 2 in the present experiment. In this condition, the orientation of the circular arc was constant and known to the subject. This knowledge could have been used by the subject to improve performance, according to our predictions derived from the pyramid model (rule 2a in the Discussion section of Experiment 2). More exactly, performance in this condition was expected to be comparable to performance for the case of a straight line target in Experiment 2 because the local uncertainty about the direction of a line, as predicted from any pair of dots, is the same for a straight line with unknown orientation and for a circular arc with known orientation. And, in fact, the performance in condition 2 in the present experiment for subjects MSG and ZP was similar to their performance in the case of a straight line target in Experiment 2.

Clearly, the results of this experiment are consistent with the predictions of the pyramid model. Only some information about the target can be used in a detection task. The information which can be used is restricted to information that can be expressed as a set of simple rules that can be applied locally by nodes at different levels of the pyramid. Thus, the information about the shape of an irregularly shaped target is not likely to be useful locally because this shape is a global property. The orientation of a known shape, however, can be used locally.

We conclude that the psychophysical results obtained in all three experiments can be adequately explained by our pyramid model, after the new definition of smoothness is supplemented by the possibility of using some prior knowledge about the local properties of a target. This knowledge can be implemented in the form of additional rules similar to rules 2, 2a or 3 in the current version of the model.

SUMMARY AND CONCLUSIONS

In this paper we have analyzed one of the Gestalt laws of organization that are involved in the phenomenon of figure–ground segregation, namely, the law of good continuation. Almost all prior theories assumed that figure–ground segregation must operate in a bottom-up fashion and, thus, is pre-attentive. This assumption seemed to be necessary to assure that this first stage of visual processing is fast. We showed, however, that this assumption is too restrictive and it does not lead to a psychologically plausible explanation of the phenomenon of figure–ground segregation.

To explain the perceptual mechanisms underlying the law of good continuation, we proposed a new theory based on an exponential pyramid algorithm. This theory is different from prior theories in several respects:

1. It does not assume purely bottom-up (pre-attentive) processing. Instead, it also allows for the effect of knowledge of, or familiarity with, the target to be detected.
2. It does not involve purely local or purely global analysis of an image. The processing involves a combination of both these types of analysis.
3. Our definition of smoothness is rooted in the likelihood, rather than the simplicity, principle.

We showed that our new model can account for a wider range of psychophysical results, as compared to prior models. Furthermore, the new predictions of our model were confirmed by the results of our experiments. We believe that this model can be further elaborated so that it can explain not only the law of good continuation, but also the general phenomenon of figure–ground segregation.

Finally, we briefly discuss one element of the methodology of our approach. Traditionally, psychological research has been conducted by first formulating a new theory that made predictions qualitatively different from predictions of prior theories [e.g. Hering's (1878), opponent process theory vs Young–von Helmholtz's (1852), trichromatic theory of color vision], and then performing experiments testing these predictions [e.g. Hurvich and Jameson's (1951), experiment on fusion of

green and red]—we call this approach "qualitative". Since the predictions were qualitatively different, the results of such experiments were usually unequivocal and thus easy to interpret. Another approach (we call it quantitative) involves formulating a new theory that made predictions different from the predictions of prior theories, but the differences did not have to be qualitative. In such cases, experiments were usually followed by mathematical or computational modeling and fitting the model to the data. The theory that led to the better fit was accepted (if none of them provided a satisfactory fit, none was accepted). The advantage of the "quantitative" approach was that it involved quantitative models and, thus, the theory was formulated in a more precise way. The disadvantage was, however, that in the presence of variability in the subject's responses it is often difficult, if even possible, to decide conclusively that one theory is clearly better than others if the only difference between the theories being compared is the degrees of relationships, rather than their presence or direction [e.g. see Luce (1986) for a discussion of different models of the speed–accuracy tradeoff].

Our approach is more conservative. On the one hand, we believe that formulating a theory in a quantitative way is important because only then can one be sure that no implicit assumptions are involved in the theory. On the other hand, we believe that the comparison of the new theory to previous theories should be made by deriving, from the new theory, predictions that are qualitatively different from the predictions of prior theories. The new theory is accepted if experimental results are *quantitatively* consistent with the new theory and *qualitatively* different from predictions of prior theories. By doing this one can avoid the problems inherent in a purely qualitative or purely quantitative approach and retain the advantages of both approaches [in fact, we applied this kind of approach not only in this present study, but also in our prior work: Pizlo (1994); Pizlo *et al.* (1995); Pizlo & Salach-Golyska (1995)].

REFERENCES

Alter, T. D. & Basri, R. (in press). Extracting salient curves from images: an analysis of the saliency network. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. IEEE Computer Society Press, in press.

Beck, J. (1966). Effect of orientation and of shape similarity on perceptual grouping. *Perception and Psychophysics, 1*, 300–302.

Ben-Av, M. B. & Sagi, D. (1995). Perceptual grouping by similarity and proximity: Experimental results can be predicted by intensity autocorrelations. *Vision Research, 35*, 853–866.

Braunstein, M. L. (1976). *Depth perception through motion*. New York: Academic Press.

Burbeck, C. A. & Yap, Y. L. (1990). Two mechanisms for localization? Evidence for seperation-dependent and separation-independent processing of position information. *Vision Research, 30*, 739–750.

Burt, P. J. (1981). Fast filter transforms for image processing. *Computer Graphics and Image Processing, 16*, 20–51.

Caelli, T. M., Preston, G. A. N. & Howell, E. R. (1978). Implications of spatial summation models for process of contour perception: A geometric perspective. *Vision Research, 18*, 723–734.

Campbell, F. W. & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology, 197*, 551–566.

Cutting, J. E. (1986). *Perception with an eye for motion*. Cambridge, MA: MIT Press.

Dobbins, A., Zucker, S. W. & Cynader, M. S. (1989). Endstopping and curvature. *Vision Research, 29*, 1371–1387.

Field, D. J., Hayes, A. & Hess, R. F. (1993). Contour integration by the human visual system: Evidence for a local "association field". *Vision Research, 33*, 173–193.

Finney, D. J. (1971). *Probit analysis*. Cambridge: Cambridge University Press.

Gibson, J. J. (1950). *The perception of the visual world*. Boston: Houghton Mifflin.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Grossberg, S. & Mingolla, E. (1985a) Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading. *Psychological Review, 92*, 173–211.

Grossberg, S. & Mingolla, E. (1985b) Neural dynamics of perceptual grouping: Textures, boundaries, and emergent segmentations. *Perception and Psychophysics, 38*, 141–171.

von Helmholtz, H. (1852). On the theory of compound colors. *Philosophical Magazine, 4*, 519–534.

Hering, E. (1878). *Zur Lehre vom Lichtsinn*. Vienna: Gerold.

Hurvich, L. M. & Jameson, D. (1951). The binocular fusion of yellow in relation to color theories. *Science, 114*, 199–202.

Johansson, G. (1977). Spatial constancy and motion in visual perception. In Epstein, W. (Ed.), *Stability and constancy in visual perception: mechanisms and processes* (pp. 375–419). New York: Wiley.

Jolion, J. M. & Rosenfeld, A. (1994). *A pyramidal framework for early vision*. The Netherlands: Kluwer, Dordrecht.

Julesz, B. (1962). Visual pattern discrimination. *IRE Transactions on Information Theory, 2*, 84–92.

Koenderink, J. J. (1984). The structure of images. *Biological Cybernetics, 50*, 363–370.

Koffka, K. (1935). *Principles of Gestalt Psychology*. New York: Harcourt Brace.

Leeper, R. (1935). A study of a neglected portion of the field of learning—the development of sensory organization. *Journal of Genetic Psychology, 46*, 41–75.

Luce, R. D. (1986). *Response times*. New York: Oxford University Press.

Macmillan, N. A. & Creelman, C. D. (1991). *Detection theory: A user's guide*. Cambridge: Cambridge University Press.

Marr, D. (1982). *Vision*. New York: W. H. Freeman.

Meer, P. (1989). Stochastic image pyramids. *Computer Vision Graphics and Image processing, 45*, 269–294.

van Oeffelen, M. P. & Vos, P. G. (1983). An algorithm for pattern description on the level of relative proximity. *Pattern Recognition, 16*, 341–348.

Pizlo, Z. (1994). A theory of shape constancy based on perspective invariants. *Vision Research, 34*, 1637–1658.

Pizlo, Z., Rosenfeld, A. & Epelboim, J. (1995). An exponential pyramid model of the time-course of size processing. *Vision Research, 35*, 1089–1107.

Pizlo, Z. & Salach-Golyska, M. (1995). 3-D shape perception. *Perception and Psychophysics, 57*, 692–714.

Rock, I. (1983). *The logic of perception*. Cambridge, MA: MIT Press.

Rosenfeld, A. (1990). Pyramid algorithms for efficient vision. In Blakemore, C. (Ed.), *Vision: Coding and efficiency* (pp. 423–430). Cambridge: Cambridge University Press.

Sha'ashua, A. & Ullman, S. (1988). Structural saliency: The detection of global salient structures using a locally connected network. In *Proceedings of the International Conference on Computer Vision* (pp. 321–327). Washington, D.C.: IEEE Computer Society Press.

Shepard, R. N. & Cooper L. A. (1982). *Mental images and their transformations*. Cambridge, MA: MIT Press.

Smits, J. T. S. & Vos, P. G. (1986). A model for the perception of curves in dot figures: The role of local salience of "virtual lines". *Biological Cybernetics, 54*, 407–416.

Smits, J. T. S., Vos, P. G. & van Oeffelen, M. P. (1986). The perception of a dotted line in a noise: A model of good continuation and some experimental results. *Spatial Vision, 1*, 163–177.

Tanimoto, S. & Pavlidis, T. (1975). A hierarchical data structure for picture processing. *Computer Graphics and Image Processing, 4*, 104–1119.

Treisman, A. & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology, 12*, 97–136.

Uttal, W. R. (1975). *An autocorrelation theory of form detection.* Hillsdale, NJ: Lawrence Erlbaum.

Vos, P. G. & Helsper, E. L. (1991). Modelling detection of linear structure in dot patterns. In Dvorak, I. & Holden, A. V. (Eds), *Mathematical approaches to brain functional diagnostics* (pp. 429–440). New York: Manchester University Press.

Watt, R. J. (1987). Scanning from coarse to fine spatial scales in the human visual system after the onset of the stimulus. *Journal of the Optical Society of America A, 4*, 2006–2021.

Wertheimer, M. (1923/1958). Principles of perceptual organization. In Beardslee, D. C. & Wertheimer, M. (Eds), *Readings in perception* (pp. 115–135). Princeton, NJ: D. van Nostrand.

Yuen, H. K., Princen, J., Illingworth, J. & Kittler, J. (1990). Comparative study of Hough transform methods for circle finding. *Image and Vision Computing, 8*, 71–77.

Zeki, S. (1993). *A vision of the brain.* Oxford: Blackwell Scientific Publications.

Zucker, S. W. (1985). Early orientation selection: Tangent fields and the dimensionality of their support. *Computer Vision, Graphics, and Image Processing, 32*, 74–103.

Zusne, L. (1970). *Visual perception of form.* New York: Academic Press.

Acknowledgement—The authors wish to thank the anonymous reviewers for useful comments and suggestions.